











Digitized by the Internet Archive  
in 2024 with funding from  
University of Alberta Library

[https://archive.org/details/ajer\\_1964](https://archive.org/details/ajer_1964)



# The Alberta Journal of Educational Research

Vol. X, No. 1

March, 1964



THE COMMITTEE ON EDUCATIONAL RESEARCH

*Faculty of Education*

*University of Alberta*





# OCCUPATIONAL ASPIRATIONS OF NORTHERN SASKATCHEWAN STUDENTS\*

WILLIAM D. KNILL  
*University of Alberta*

## Introduction

One of the major steps every young person takes is going to work. There may be little or no actual choice on his part. His economic status may decide how he will earn a living; his society may limit his choice of occupation through racial discrimination. His intellectual ability, level of education and training, and his physical attributes—all such factors will influence his job history and work prospects.

What aspirations and expectations do young Northern students disclose when asked, "What do you want to be when you grow up?"

Kurt Lewin wrote as follows about levels of aspiration.

A successful individual typically sets his next goal somewhat, but not too much above his last achievement. In this way he steadily raises his level of aspiration. Although in the long-run he is guided by his ideal goal, which may be rather high, nevertheless his real goal for the next step is kept realistically close to his present position. The unsuccessful individual, on the other hand, tends to show one of two reactions: he sets his goal very low, frequently below his past achievement—that is, he becomes intimidated and gives up reaching out toward higher goals—or he sets his goal far above his ability. This latter conduct is rather common. Sometimes the result is a gesturelike keeping up of high goals without serious striving; it may at other times mean that the individual is following blindly his ideal goal, losing sight of what in the present situation is possible. To develop and to maintain high goals and, at the same time, to keep the plan for the next action realistically within the limits of what is possible, seems to be one of the basic objectives for a criterion of high morale.<sup>1</sup>

Another writer, F. La Violette states that each member of a minority, as part of his struggle for identity, must make a crucial effort to control his own fate. This struggle—

Will be marked by the hopes of younger people for new jobs, for places in the professions, or for better conditions in logging and in the catching and packing of fish. Questions of schools, scholarships, and undeveloped abilities are peppered through the records of the Joint Committee . . . This struggle for identity involves a conception of educational opportunities and satisfying occupational choices.<sup>2</sup>

## The Research Design

*Data Collection.* Letters were sent out to all teachers or principals of Northern Saskatchewan schools who had students in

\*This study was made financially possible by The Centre for Community Studies, University of Saskatchewan and the cooperation of Dr. A. K. Davis, Senior Research Sociologist. Miss Muriel Case assisted in data tabulation. However, the writer alone accepts the responsibility for all interpretations made and conclusions drawn.

1 In Krech and Crutchfield, *Theory and Problems in Social Psychology*, N.Y., 1948, p. 410.

2 *The Struggle for Survival*, Toronto, p. 185.



Grade VII and above. This included Creighton and Island Falls, but not Uranium City. The teachers were asked to have each student write an essay on the subject, "What I want to be when I grow up." The administration of the assignment was spelled out briefly to emphasize the value of getting a spontaneous response from the child. The letter said in part:

... the children's own, free, undirected responses, not anything suggested or prompted by the teacher. The conventional essay, saying all the 'right' things, written in painstaking best English, would be almost valueless for our purposes.

A suggestion was made that the essay should take from 30 to 60 minutes for completion. The teachers were assured that the essays would not be judged for academic achievement or for evaluating teacher competence.

Each essay was to be accompanied by additional information: (a) the name of the school, (b) the student's name, (c) grade, (d) age, (e) sex, and (f) ethnicity.

*The Sample Population.* The project aimed for a total population of all school children in Grades VII and above of the selected schools. With the co-operation of the Northern teachers, 266 essays were received—representing 71 percent of the senior elementary students in school in the Northern Administrative Areas, excluding Uranium City and Island Falls. Seven essays from the Gunnar school were rejected because of inadequate identifying information regarding age, grade and sex. We believe we obtained an adequate cross-section of senior students in Northern Saskatchewan schools.

*Data Analysis.* A content analysis of the essays was undertaken. No previous studies of this nature have been conducted with minority groups, to the writer's knowledge. Therefore, it was necessary to construct our own system of analysis. The study proceeded within the definition of content analysis set by Berelson: "... a research technique for the objective, systematic, and quantitative description of the manifest content of communication."<sup>3</sup>

The essays were read through first to establish the general themes which appear. An analysis sheet, in the form of a check-list, was then constructed, and the essays were re-read carefully in random order. At the conclusion of this step, six more categories were added, and the essays were read for the third time. The check-lists were then coded and processed on punched cards for IBM equipment. The coding permits quick references to any single essay, thus facilitating the location of appropriate quotations illustrating various points used below. In the final preparation of this report, the essays were read in their entirety for the fourth time. This

---

3 *Content Analysis in Communication Research*, Glencoe, 1952, p. 18.



was an attempt to keep in sight the unique content and character of the essays, in order to avoid becoming completely involved with sheer quantification.

### The Findings

Judged by the great discrepancy between our respondents' occupational preferences and the actual occupational distribution of the employed Northern population, Northern students have somewhat unrealistic expectations regarding their future jobs. If a Northland community were solely inhabited by these students, each in the job of his choice, we would have 170 professional people. There would be one doctor for every 45 people, one nurse for every two people, nine civil engineers and 34 teachers. Two storekeepers would offer competition to each other, and two lawyers would be available for legal contests. Transportation would be a major

TABLE I

COMBINED FIRST THREE OCCUPATIONAL CHOICES OF  
266 STUDENTS ATTENDING NORTHERN SCHOOLS—BY SEX  
(Grades VII-X, inclusive)

Occupational Category	Occupational Choices					
	Boys		Girls		Total	
	No.	Percent	No.	Percent	No.	Percent
(1) Managerial .....	2		0		2	
Professional .....	44		112		156	
Theological .....	2		8		10	
Sub-total .....	48	32	120	64	168	50
(2) Clerical .....	0		17		17	
Commercial .....	0		6		6	
Service .....	0		17		17	
Sub-total .....	0	0	40	21	40	12
(3) Transportation .....	35	23	15	8	51	15
(4) Agriculture .....	4		1		5	
Fishing-Trapping .....	7		1		8	
Sub-total .....	11	7	2	1	13	4
(5) Mining .....	5		0		5	
Manufacturing-						
Mechanics .....	20		1		21	
Sub-total .....	25	17	1	1	26	8
(6) Armed Services .....	6		0		6	
Law Enforcement .....	14		3		17	
Sub-total .....	20	13	3	2	23	7
(7) Professional Sport .....	9		0		9	
Arts .....	1		5		6	
Sub-total .....	10	7	5	3	15	4
GRAND TOTAL .....	149	99	186	100	336	100



industry, with 140 people employed to transport the members of the community, mostly by air. The present-day basis of the Northern economy, namely fishing, trapping, and mining would be represented by six fishermen-trappers, and five geologist-prospectors. Four conservation officers, three forest rangers and a police of four would constitute the law and policy enforcement body of the community. Five professional hockey and baseball players and one professional singer would provide the community's entertainment. There would be no caretakers, community maintenance workers, no guides, fish packers, or lumber jacks.

From the foregoing, we can surmise that many students must revise their choices. They will need information about occupations which they are not seriously considering at the present time, and also they will require information about opportunities for employment, both in the North and "outside."

*Occupational Choices of Boys and Girls:* We shall now turn directly to the data. Table I presents the occupational choices classified by sex. If a student gave a second or third choice of an occupation, it is included in this tabulation.

The usual, obvious sex differences in occupational choice appear. Indeed, Table I is an illustration of occupational sex-typing in North America society. Girls choosing professional careers are chiefly interested in nursing and teaching; the boys are chiefly interested in engineering, teaching, medicine, and research—in that order. In the second group of occupations, the girls wish to be stenographers, hairdressers, and store clerks; but no boys chose clerical, commercial or service careers. Over one-fifth of the boys chose careers in transportation—most of these wish to be airplane pilots (bush, airline, and jet). The girls in this category wish to be stewardesses. In the fourth group, the boys favor mechanics over fishing, trapping and farming. The one girl who wants to be a car mechanic confides, "That has been my ambition ever since the day I first stepped into a garage . . . The laughing kids that think I'm crazy will not change my mind. I may be a girl, but girls can be just as good as boys." Twenty boys chose the armed services and law enforcement agencies, and nine chose professional sports. As many boys want to be professional hockey players as R.C.M.P. officers.

Apparently the students are choosing occupations within their field of vicarious experiences. They aim at those roles with relatively high community status, and, in some cases, with an aura of glamour. Some base their choice upon information gleaned from friends, books, and motion pictures. In 114 essays (42.9 percent), specific reference was made to the need for further education to attain their goals. It is noteworthy that not a single student aspired



to be a Hollywood or TV star, a millionaire, or rodeo cowboy. Only one thought he would like to be an astronaut. In this respect they are not unrealistic, and have moved away from childish ambitions.

The single student who aspires to outer space also sees himself in an organization which can offer too much routine and security. He is prepared to pay something for some "life-taking risks." Thus states a 14-year-old boy in Grade VIII:

If I am in the Airforce I would like to be one of the first men to the moon or test pilot because I don't want a dull life which all you do is go to the plane, check it, take off, land and leave it. I want something exciting which consists of continuous danger, hard work, and my two-bits worth of life-risking tasks."

Two general impressions are gained from the essays.

- (1) The majority of students are selecting occupations which require educational facilities not available to the Northern student. One 15-year-old Métis student in Grade X states, ". . . Many here in Cumberland have wished to stay on in school but have quit after finishing Grade IX. The only reason probably being that there is no high school here and none close enough to go to." This suggests that few of these students will attain their objectives. Few of the Métis-Indian girls will become nurses, teachers, stenographers, or stewardesses. The boys will have a better chance, but not many become doctors, engineers, pilots, or research scientists under present conditions.
- (2) The students are selecting careers on a basis of little knowledge of the training or educational requirements for the career, and very little information about the actual job requirements. A typical statement is made by a 16-year-old Métis student in Grade X. "Every since I can remember, I've wanted to be a nurse. Maybe it was the starchy white uniforms I've seen them wear or maybe it was just seeing what they do to help people. The reason for my decision is because I would like to see the world, since the farthest south I've been is Prince Albert."

*Comparative Occupational Choices of White and Métis-Indian Girls.* Table II shows the combined first three occupational choices of the girls who wrote essays for us. Over half of the choices are for nursing and teaching. No ethnic differences appear in the preferences for those two occupations.

Those occupations which attract mainly White girls are clearly higher in status, income and educational requirements than those which attract Métis-Indian girls. But let us underline the fact that many Métis-Indian girls want to be teachers and nurses. This shows an awareness of, and orientation toward, upward mobility in White society. White girls display more awareness, and higher

TABLE II  
OCCUPATIONAL CHOICES OF 146 NORTHERN GIRL  
STUDENTS—BY ETHNICITY

	White		Métis-Indian		Total	
	First Three Choices	First Choice	First Three Choices	First Choice	Combined Choice	Percent
Nurses .....	36	27	34	31	70	35.0
Teachers .....	19	11	17	14	36	18.0
Office Stenos .....	11	5	6	6	17	8.5
Airline Stewardesses...	8	5	5	4	13	6.5
Hairdressers, Nurses Aides .....	4	2	7	6	11	5.5
Missionaries, Nuns .....	2	2	5	2	7	3.5
Store Clerks .....	2	2	4	4	6	3.0
Lab Technicians, Dental Assistants...	5	3	1	1	6	3.0
Doctors, Scientists, Social Workers .....	5	4	1	1	6	3.0
Artists .....	1	1	2	0	3	1.5
Policewomen .....	1	1	2	2	3	1.5
Entertainers .....	1	0	1	0	2	1.0
Pilots .....	0	0	2	2	2	1.0
Farmers .....	1	1	0	0	1	.5
Trappers .....	0	0	1	1	1	.5
Mechanics .....	1	1	0	0	1	.5
Others .....	8	5	6	2	14	7.0
	105	70	94	75	199	99.5
Number of Respondents .....	70		76		146	

In what occupational preferences do ethnic differences show up? White girls clearly outnumber Métis-Indian girls in aspiring to the following occupations:

- doctors, scientists, social works
- lab technicians, dental assistants
- airline stewardesses
- office stenos

On the other hand, Métis-Indian girls lean more strongly than White girls to certain other occupations:

- missionaries, nuns
- store clerks
- hairdressers, nurses aides

aspirations. But Table I indicates that a majority of Métis-Indian girls are clearly looking upward toward White middle-class occupations for their future.

Religious callings appeal more strongly to Métis-Indian girls. The two groups aspire in equal proportions only to nursing and teaching—these professions, however, draw by far the largest number of preferences. Other occupational choices are too sparsely scattered among various types of work to permit any ethnic comparisons.



Thus far, we have been discussing the combined first three occupational choices. Because it is not always clear in the essays which is the first choice, or whether the indicated first choice reflects a definite inclination on the part of the student—it has seemed best to base our analysis on the combined preferences. Not the least important advantage this affords is the larger number of cases. However, this decision is debatable. We have therefore shown the first choice alone in Table II. How do first choices differ from the combined choices?

The list of occupations wherein differences in preferences appear along ethnic lines becomes shorter. White girls choose the following occupations more frequently than do Métis-Indian girls:

- doctors: scientists, social workers
- lab technicians, dental assistants

Métis-Indian girls are predominant among those preferring to become—

- store clerks
- hairstylists, nurses aides

In Canadian society it is commonplace for young people at one time or another to want to be doctors when they grow up. Is it not surprising then, that—as yet—not one of our 76 Métis-Indian girl students say she wants to be a doctor? Is there any implication for the public image of social workers among the Northern Métis and Indian people that not one Métis or Indian girl expresses any desire to be a social worker? Only one Métis girl says she would like to be a scientist. Even the world-be lab technicians are all White girls. The future nurses aides, on the other hand, are all Métis-Indians. We believe that these occupational preferences reflect all too clearly the class cleavage—and the corresponding differences in life prospects—between the two ethnic groups. Aspirations toward higher occupational status—in the White economic world—have emerged among Métis-Indian young people, but they have not yet developed very far.

*Occupational Preferences of White and Métis-Indian Boys.* Table III shows the combined first, second and third choices of 71 White and 47 Métis-Indian boys enrolled in Grades VII-X in Northern Saskatchewan schools. The column of totals is arranged in order of descending numerical importance.

As previously indicated, the scientific professions rank as first choice for the group as a whole, with engineers and pilots second, and skilled workers third. Together, these three occupational categories account for half of the male students' preferences.

Only in the last of these three occupational categories, however, are the preferences of White and Métis-Indian students proportion-

TABLE III  
COMBINED FIRST THREE OCCUPATIONAL CHOICES OF  
177 NORTHERN BOYS—BY ETHNICITY

	White		Métis-Indian		Total	
	No.	Percent	No.	Percent	No.	Percent
Doctors, Engineers, Scientists, Geologists, Social Workers .....	27	26.5	10	14.1	37	21.4
Railroad engineers, Air pilots .....	23	22.6	8	11.3	31	17.9
Skilled workers (car- penters, mechanics, electricians .....	11	10.8	9	12.7	20	11.6
Teachers .....	4	3.9	7	9.9	11	6.4
Ball players, hockey players .....	4	3.9	5	7.0	0	5.2
Conservation officers, Forest rangers .....	5	4.9	2	2.8	7	4.1
R.C.M.P. ....	5	4.9	2	2.8	7	4.1
Trappers, Fishermen .....	3	2.9	4	5.6	7	4.1
Armed Services .....	2	2.0	4	5.6	6	3.5
Truck and bus drivers ..	0	.....	4	5.6	4	2.3
Farmers, ranchers .....	3	2.9	1	1.4	4	2.3
Storekeepers .....	0	.....	2	2.8	2	1.2
Priests .....	0	.....	2	2.8	2	1.2
Others .....	15	14.7	11	15.5	26	15.0
Totals .....	102	100.0	71	99.9	173	100.3
Number of respondents	70		47		117	

ately balanced. The professions and the pilots-engineers attract notably more White than Métis-Indian boys.

Let us now list separately those occupational choices which differentiate the two ethnic groups. For most of these categories, the N (number of cases) is too small to warrant any claim of statistical significance.

*Occupations Attracting Proportionately More White than Métis-Indian Boys.*

- the scientific professions
- railroad engineers, air pilots
- conservation officers, forest rangers
- R.C.M.P. officers

*Occupations Attracting Proportionately More Métis-Indian than White Boys.*

- teachers
- baseball and hockey players
- truck and bus drivers

If we consider simply the first occupational choices of the 117 Northern schoolboys, we find approximately the same pattern of similarities and differences in the occupational preferences of the two ethnic groups that we have described above with respect to



the combined three choices. Some conclusions are suggested by the job choices expressed by the boys.

Among the boys, *first*, there are obvious ethnic differences in certain occupational preferences, although the *range* of choices is nearly the same for both ethnic groups. Larger proportions of White than of Métis-Indian boys aspire to the high-status scientific professions, and to such glamorous jobs as railway engineers and airplane pilots. Only three occupations are monopolized by one ethnic group—truck drivers, storekeepers, and priests—all three by the Métis-Indians.

*Second*, the level of Métis-Indian job aspiration seems somewhat lower than that of the White boys. This may indicate realism on the part of the boys—given the obvious ethnic differences in life prospects. Kew writes of Cumberland House:

The most significant feature about employment inside the community is the unequal sharing of opportunities between the two main ethnic groups—White and Métis-Indians. All the White wage-earners, but only a fraction of the Métis-Indians are employed within the community. Furthermore the White occupy all the superior positions. Most Métis work under White superiors, and no White under Métis. Finally, those few Métis who do have local permanent jobs do not earn as much as the Whites.<sup>4</sup>

However, our student essayists do not see this class structure explicitly as “discrimination.” When the Métis student refers to his parents’ occupational status as too low for his aspirations, he generally attributes this to his parents’ lack of education. The Métis child perceives education as the great “equalizer”—the avenue to equal job opportunity. One Grade IX student describes his hopes:

For my parents, the pay they get is good enough because that’s about what they’ll be getting all their lives. It is very different with me. I want to finish school, get a good job with good pay, and some day have a family of my own who will think the same as I. I don’t blame my parents for not having a good education—they did not have the chance we have.

*Third*, certain individual items then claim our special notice.

- none of our 47 Métis-Indian schoolboy respondents wanted to be a social worker or lawyer;
- only two of the 47 wanted to be priests;
- not one Métis-Indian boy wanted to be a Conservation Officer, either as first, second or third choice! In view of the unquestionable importance, power, and prestige of conservation officers in Northern Saskatchewan, we think this stark fact should have a sobering impact on many readers. Does it mean that Métis-Indian boys see the Conservation Officer as too high and remote for their aspiration? Or do they reject him as an alien power figure? Or is there some other explanation? We cannot say.

<sup>4</sup> *Cumberland House in 1960*, p. 61 (Economic and Social Survey of Northern Saskatchewan, Centre for Community Studies, Saskatoon, 1962.)

## REASONS FOR OCCUPATIONAL CHOICE, BY ETHNICITY

A 16-year-old Métis student in Grade IX notes a new outlook in his community:

Beauval today is much different from what it was ten years ago because now it seem people are starting to realize the great need for education. Before, education was not so important; but today it seems that nothing can be done without education. Today, education is most important if a person intends to do something and be a somebody.

"To be a *somebody* . . .' Seeking an identity through an occupation is a common reason for selecting certain careers. This and other reasons are illustrated in Table IV, which classified the reasons for occupational preference by the two ethnic groups. The mean number of reasons for all students is 2.7; groups were equally verbal in giving reasons for their choices.

One Métis student in Grade IX compresses into a few lines many problems of Northern education—lack of educational facilities, students dropping out of school, delinquency and racial barriers. He says:

I would like to become a teacher because I would like to help the people of the North. I know how hard it is here without our education. In most of the towns around here, the pupils quit at 15 and even sooner. Some do not get a chance to get their education. They run around town get into trouble and mischief, making gangs and fighting. I was once in a gang and I know how it feels. Also I would want more Métis to graduate and have good jobs and be somebody in life.

An analysis of the reasons given by the students for their particular career choices points up some differences which may be related to ethnic group membership. A checklist of 16 reasons was drawn up and applied to each essay. A rigorous requirement was that the essay must have clearly stated reasons before they could be counted; the analyst was not to "read into" the essays any meaning which may or may not have been intended by the student author. Table IV shows the checklist and the results.

Compared to the White respondents, the Métis-Indian boys seem to express stronger social motives, more family influences, a stronger orientation to status in the North rather than to status in general, and—surely a key fact—dislike for their parental occupations. Since the latter would be mainly trapping and fishing, we can infer a clear aversion, on the part of many Métis-Indian boys, to the traditional Indian occupations. The oncoming generation is not content to be trappers and fishermen. Northern Métis and Indian young people are definitely orienting to the modern world of White urban-industrial society.

The chief reason given for career preference was the desire to help others. This "social orientation" shows up also in the numerous students wishing to be nurses and teachers. However, students choosing such careers as bush pilots and geologists likewise see



TABLE IV  
REASONS FOR WHITE AND MÉTIS-INDIAN STUDENTS  
FOR OCCUPATIONAL CHOICE

Reasons	White (N=140)		Métis-Indian (N=126)		Total (N=266)	
	No.	Percent*	No.	Percent	No.	Percent
Social orientation: desire to help others, contribute to humanity .....	49	35.0	59	47.7	108	40.6
A job providing oppor- tunity to meet new people, hear different languages .....	64	45.7	41	32.5	105	39.8
Money-oriented .....	44	31.4	37	29.4	81	30.5
Cites a childhood experience .....	28	20.0	41	32.5	69	25.9
Wants excitement, thrills .....	42	30.0	21	18.3	63	23.7
A career with future status .....	37	26.4	25	21.1	62	23.3
Cites some family influence .....	22	15.7	26	23.9	48	18.1
Chooses a job because it would be "fun," or "I would like it." .....	17	12.1	12	9.2	29	10.9
Security-oriented .....	15	10.7	12	9.2	27	10.2
Chooses a job for its glamour .....	11	7.9	10	5.5	21	7.9
Wants high status for Northern community	11	7.9	15	13.7	26	9.8
Wants high status "outside" .....	9	6.4	4	3.7	13	4.9
Dislike for parental occupation .....	1	.7	11	10.1	12	4.5
Religious influences .....	2	1.4	5	4.6	7	2.6
States a choice but qualifies with, "I don't know," or "I'm not sure." .....	16	11.4	15	11.0	31	11.7
Average responses per student .....	2.6		2.7		2.6	

(\*percent do not add to 100, because most students gave more than one answer.)

Reasons given more  
often by White students  
—meet new people  
—excitement, thrills  
—career with future status

Reason cited more  
often by Métis-Indians  
—social orientation  
—childhood experience  
—family influence  
—desire for high status in  
  Northern community  
—dislike for parental  
  occupation

their work as directly helping their fellow man in the North. Table IV indicates a significantly greater number of Métis and Indian than of White students expressing social concerns in relation to their future jobs. This finding is consistent with the generally accepted view that Indian people have a greater concern for mutual welfare than do White people.

The "social conscience" is sometimes difficult to express, but appears in various forms, such as the diverse pattern of occupations suggested by a Grade VIII girl:

Still I would like to be a Nun. They also help people all over the world in many ways. I would want to go to Africa or China to teach the people about God and to help those suffering from diseases. After about five years of serving I would like to leave, if I couldn't go on, or if I couldn't be loyal to God, and get married or try to get another job somewhere. These two things I would like to be right now, but being a teacher or a nurse or something is still in my mind. There are so many wonderful jobs to choose from. Right now, I am not sure what my choice may be. The future will tell.

The White student is more concerned with a job that will provide him with the "opportunity to meet new people, visit foreign lands, and hear different languages." This adventurous spirit also appears again, and predominately, amongst the White students in the form of a desire for "excitement and thrills." These White students explain their job preference on such grounds as— they would "like it," or "it might be fun." We may infer that the White students are more daring, more audacious in their job speculation than are the Métis or Indian students.

Some of these students get excited even about such occupations as autobody repairman. A Grade VIII boy writes:

I would also like to be a bodyman. It seems exciting to pound out cars. From the start when you bring in a smashed car to the end when you bring it out of the garage all painted and the dents taken out of it.

One White student from Creighton expresses the modern man's dilemma: how to combine status, stability and security with a job on the *qui vive*. One White student who hopes "to live it up" on the West Coast first says:

When I leave Vancouver I will come back to Flin Flon and settle down here . . . I would like to have my own house in the uptown area not too far from Main Street. I would become a member of the Canadian Legion. I would also belong to the Salvation Army Church, I would attend Church every Sunday at 7:00. In my spare time I would build a boat and go fishing on the many lakes around Flin Flon. I would have a car of my own so I would not have to ride a taxi. When I become an engineer I hope to make a fortune.

The Métis students show a greater "family orientation" in job selection: they mention family influences and childhood experiences, usually in a family context, as reasons for preferring a certain occupation. Ten percent of the Métis students—but less than 1 percent of the Whites—admit to a definite dislike for their father's



occupation (in every case fishing or trapping). One 14-year-old Métis boy expresses this: "Maybe I won't get an opportunity to get an education but I'll try hard to get one so I don't have to make a living by fishing and trapping." A second student, 15 years old, has this to say:

I would not like to work where my father is working because I think a mechanics is much better paid and has a better chance to earn a living. My goal for the future is very high but I will try my best. Success is a very small word but has a big meaning—and that is what I will be aiming for.

The only White, a Grade VII student, to express a dislike for his father's occupation makes a very moving statement. High regard for education as a means to occupational success is obvious. This Grade VII student states:

I would not like to follow my dad's tracks, I am sorry to say. But he had never a chance to go to school past Grade 7. So down he went as a miner breaking his foot 3 times and also breaking our family's heart in sorrow. But for those that got only Grade 7 or 8 education and quit for a little job that will not last them a lifetime, you might as well push them off the road and don't bother them, ten chances to one in later years they will be leading a gang or in it. But I would like to be a person to stop them, and maybe they may try their luck at night school if they have any brains at all.

Another student feels quite the opposite—he identifies completely with his father. This White youngster also has the touch of the romantic in him as he rhapsodizes about the North:

My father's footsteps are adventurous ones and I would follow his tracks if it would be 70 years from now.

I would like to spend most of my time up North around Husdon's Bay in Ontario where the shield still lingers and the muskeg swamps are frequent enough. For there a few years from now the wilderness will still be the same and the wolves will still be chasing the deer and moose through the hills.

Occupations were chosen, in several instances, because the career offers a high status role. The Métis student tend to think in terms of higher status in the Northern community, while the White students prefer careers offering high status in "outside" communities, on a provincial or national level. It would appear that the White students have a wider horizon and greater awareness of opportunities "outside."

Consistent with this generalization is the fact that more White students see a job offering high status in the future rather than one conferring immediate status elevation. White students seem more "future oriented" than Métis Indian students. They can look more easily and realistically to long-run success than to immediate promotions. However, some Métis students look toward the next generation. Two Métis girls express their hopes for marriage:

When I get married I'm going to get a good-sized house, not too big and not too small. When my children grow up I will send them to school so

they will have a good education. I don't want them to run around town all day doing nothing. I want them to get good jobs and not be like their grandparents. I don't want them to get into trouble . . .

The second writes:

If I get married I hope I won't be poor, but I wish I'll have all the things I need like electricity, running water, water pipes, and all sorts of electric appliances, etc. When I have children I'll make them pass their Grade 12, then go to University, then be someone special.

Two Métis girls revealed a compact between themselves agreed to earlier which, although unrealistic, expresses what they perceive as a need for their community. The first says:

Ever since I was knee-high to a grasshopper I've always wanted to be a Police woman. And I would like to come in to La Ronge and work as a Police Woman.

Her associate, two years older, gives their reason for this career:

I want to come to La Ronge as a Policewoman after we take our courses. So there'll be peace in La Ronge and we can put the bad people in jail.

There appears to be little difference amongst the ethnic groups regarding the desire for jobs because of the financial remuneration. The third most common reason given was "money," equally preferred by both Métis and White. Jobs which offer security are stressed equally by both groups, although this aspect of employment did not appear to be a great concern to many. However, one Métis refers to his fear of unemployment:

When I see children running around at night in the street, I feel sorry for them and their parents. I also feel sorry for my people who stay in their houses and are out of jobs. When I see my big brothers bum around my mother, I feel sorry for them so that is why I want a future.

Education was first established by religious orders in the North, and the religious influence on education is still considered to be strong. However, it seems to have little bearing on students' occupational choices. Only seven students referred to the religious influence of their church as affecting their choices.

Political action as a means of improving the lot of Northern people was expressed only twice in the essays. One student feels that living north of Prince Albert should not deter him from becoming Prime Minister of Canada.

The lone White student who wants to be a trapper puts his case clearly, and in one way points up the characteristic rebelliousness of youth:

When you are a trapper you do not have to run by the whistle or be bossed around. If my parents would let me quit school and do what I wanted, I would head for the bush so quick you wouldn't even know I existed. I think that if you want a good career just head North and trap.

A few students were just not up to the task of choosing a career. One student begged off by declaring, "It gets complicated just thinking of the things I could be. So I just don't know."



# AN EVALUATION OF SELECTED INTELLIGENCE TESTS FOR TWO SAMPLES OF METIS AND INDIAN CHILDREN

L. W. WEST

*Edmonton Public School Board*

AND

R. S. MACARTHUR

*University of Alberta*

## **The Problem**

Contemporary philosophies of education expect that instruction will be adapted to the intellectual potential of the pupil. However, in order to adjust teaching treatments to pupil potential, a fair estimate of that potential is required. Many studies in recent years indicate that conventional verbal measures of intelligence tend to discriminate against children who are not members of urban middle-class culture. Teachers and counsellors frequently require a less biased estimate of the potential of a pupil who is the product of a rather different culture. Hence the search for measures of intelligence which are relatively independent of specific environmental experiences becomes especially important.

Several so-called culture-reduced tests have been developed in an attempt to minimize the effects of cultural differences. The present paper reports a study of the relative extent of bias evident in a selection of such tests for a sample of Faust, Alberta, Métis children and for another sample of Fort Simpson, N.W.T., Métis and Indian children.

## **Theoretical Foundation**

It is recognized that nature and nurture are inseparable in the production of intelligent behavior. In agreement with Hebb(4), it is postulated that individual differences in measured intelligence result from the interaction of heredity and environment in a multiplicative manner. Moreover much of the confusion and controversy over the relative importance of these factors has resulted from varied uses of the term "intelligence." For this reason Hebb distinguishes between intelligence A (innate intellectual potential) and intelligence B (present level of performance). Recognizing that different tests will assign different values to an individual's present level of performance, Vernon(8) added intelligence C (an estimate of B made by a given intelligence test).

It is suggested here that if the concept of *potential* in contradistinction to proficiency is to have any place in psychological theory it is necessary to further distinguish between intelligence

A (i.e. *innate potential*) and *present potential*. Present potential may be defined as an individual's present capacity for future development of intelligent behavior. For convenience we may refer to present potential as intelligence A'. A' is the residual of an individual's original innate potentiality at any particular time after conception.

Whereas intelligence A is a constant, intelligence A' is a variable such that A' is less than or equal to A. Such factors as nutrition, toxic influences, destructive or traumatic events, and absence of stimulating experiences may intervene to cause intelligence A' to be less than intelligence A. The essential difference between intelligence A' and intelligence B at any time is that A' refers to the potential or capacity for future development assuming novel and optimum treatment. Intelligence B, on the other hand, refers merely to the level of presently developed and functioning intellectual abilities.

The assumption that intelligence B results from the interaction of innate pre-dispositions and environmental influences in a multiplicative manner gives rise to certain interesting logico-mathematical deductions. First, the assumption may be expressed by the mathematical model  $B=AE$  where:

B denotes intelligence B as defined by Hebb,

A denotes intelligence A, and

E represents the totality of cultural or environmental experiences, which stimulate, thwart and direct organized behavioral development.

From this paradigm it becomes obvious that attempts to hold the effects of E constant are tantamount to attempts to restrict the variability of B to that of A and thus provide an estimate of innate potential. The magnitude of A obviously can never be measured directly but can only be inferred from the magnitude of B. When E is set equal to zero by simple algebraic substitution in the equation  $B=AE$ , B also becomes zero. This fact implies that it is impossible to design a test of B which is not contingent upon E.

However impossible it may be to design a test which is culture-free, the task of designing culture-reduced tests need not be abandoned. The culture-reduced tests represents an attempt to minimize the influence of: (a) specific information and skills that not all children have had equal opportunity to acquire, and (b) personality factors more likely to be developed in one culture than in another, e.g. motivation and work habits. It is suggested that by minimizing the variance in intelligence B attributable to these environmental factors, the resultant measure will be an estimate not of intelligence A but of A'. Although it appears futile at present to attempt even an indirect estimate of intelligence A, an assessment of A' is a realistic goal. Further, it is present potential in contradistinction to



innate potential that it of interest and practical significance to the educator. It is present potential on which future development depends, and upon which adaptive teaching treatments must be based.

Hypotheses

The essential purpose of the study reported in this paper was to answer an urgent and practical question: "Which of the available group tests or sub-tests assess intellectual ability with a minimum of cultural bias for samples of Métis and Indian children?" The following hypotheses were thus set up:

- I. Some of the tests in the experimental battery will show considerably less cultural bias than will others.
- II. The tests which show least cultural bias will consists largely of items that can be solved in any language or mode of expression and which are probably as novel to one group as to another.<sup>1</sup>
- III. Tests which show least cultural bias will, nevertheless, show substantial correlation with school achievement.
- IV. Tests which show least cultural bias will show substantial correlation with other tests of the battery.

Experimental Design and Procedure

*Samples:* Data used in the study were collected for two samples of Métis and Indian children. The Faust sample consisted of all Métis children in grades I to VIII inclusive who were attending the Faust School in May, 1961, and for whom complete test data were available. This sample consisted of a total of 126 pupils.

The Fort Simpson sample, which permitted a complete replication of the investigation, consisted of all Métis and Indian children in grades I to IX inclusive, who were attending the Fort Simpson School, at Fort Simpson, N.W.T. in October, 1961, and for whom complete test data were available. This sample consisted of 155 pupils.

The two samples were sub-divided for analysis by grade level, as follows:

Grade I	Faust N = 32	Fort Simpson N = 19
Grade II and III	N = 42	N = 46
Grade V and VI	N = 29	N = 58
Grade VII and VIII	N = 23	N = 32

*Measures:* The following battery of tests, selected by MacArthur for a factor analytic study, provide the instruments for analysis in this study: Progressive Matrices (Standard and Coloured), IPAT Cattell Test of "g" Scale 2, Lorge-Thorndike Non-Verbal Intelligence Test (Levels 1, 2, 3, 4, with Part I, Part II, Part III and Total scores for each level), Safran Culture-Reduced Intelligence

<sup>1</sup> In order to test this hypothesis each of the tests was inspected and categorized into two groups prior to the analysis of data. Tests which appeared to fit the above description were classified as "culture-reduced." Those which did not were classified as "conventional."

Test, California Short-Form Test of Mental Maturity (Elementary, Primary, Pre-Primary, with six sub-tests and a Total score at each level), and Detroit Beginning First-Grade Intelligence Test.

*Administration:* At Faust, the battery of Tests was administered in May 1961 by home room teachers, under the supervision of Mac-Arthur. At Fort Simpson, the battery was similarly administered in October 1961.

*Method of Testing Hypotheses:* Raw scores on each test for the two samples at the grades I to VI levels were converted to derived scores based on a scale of Calgary T-scores, the Calgary group having a mean of 50 and standard deviation of 10 on each test. For the Grade VII and VIII level, raw scores were converted to derived scores based on a scale of Edmonton T-scores. This procedure made it possible to compare the performance of the samples on the different tests of the experimental battery. Differences in performance on the several tests relative to that of the urban white standardization group were interpreted as indicative of the extent of cultural bias in the tests.

The mean, standard deviation, mean correlation with other variables of the battery, and correlation with school achievement were found for all tests. For these computations, the I.B.M. 1620 was used.

Hypothesis I was then tested by submitting all observed differences between means for each of the grade levels to the appropriate statistical test for significance (Duncan's multiple-range test, or the ordinary t-test).

For testing Hypothesis II, scores for all tests in each classification were combined and the mean and standard deviation of the combined groups were found. The difference between the means of the two groups was then submitted to the traditional t-test. A one-tailed test was used since the direction of the difference was hypothesized. Performance which was significantly higher on one group of tests than on the other was again interpreted as indicative of less cultural bias in the former group of tests.

Hypothesis III was tested by finding the correlation between all variables at each level and grade placement on the California achievement battery for that level. Hypothesis IV was tested by finding the average of all non-spurious correlations of a given test with every other variable of the battery, using Fisher's z transformation.

*Criteria for Evaluation:* Four criteria were used to evaluate each test as a cross-cultural measure of intellectual potential. These



four criteria, generally speaking, parallel the four hypotheses previously set up. They are as follows:

1. A cross-cultural test of intellectual potential should show less difference between cultures in a bi-cultural administration than do conventional verbal tests.
2. It should contain items that can be solved in any language or mode of expression and which are likely to be as familiar and useful for one cultural group as another.
3. It should show substantial relationship to school achievement.
4. It should show a significant relationship with other well known and commonly used measures of intelligence.

The results from testing Hypotheses I to IV suggest that the degree to which a given test meets the evaluation criteria is a function of the level at which the test is administered. In other words, a test may satisfy the criteria at one level, but may fail to do so at another. For this reason the tests have been evaluated at four different grade levels.

Since the major purpose of the study was to identify those tests which showed a minimum of cultural bias, not all criteria were considered of equal importance. Each test was therefore evaluated primarily against criterion 1, provided that it met the minimum requirements of criteria 2, 3, and 4. A test was regarded as having met criterion 2 if it had been classified a priori as culture-reduced. It was regarded as having met criterion 3 if it showed a correlation of .40 with the California Achievement Total grade placement, provided however, that such a correlation was significant at the .05 level. Finally a test was regarded as having met criterion 4 if it showed a mean non-spurious correlation with other tests of the battery which is significant at the .05 level.

## Results

The main results from this study are summarized in Tables I to IV. For both the Faust and the Fort Simpson samples, and at each of the four grade levels, some of the tests of the experimental battery were found to show significantly less cultural bias than others.<sup>2</sup> These test instruments have been identified and Hypothesis I has received support.

In each instance where it was possible to test Hypothesis II it was found that tests which (a) consist largely of items that can be solved in any language or mode of expression, (b) have minimal dependence on past specific learning, and (c) are probably as novel to one culture group as to another, showed significantly less cultural bias than other more conventional tests. Thus Hypothesis II was supported.

<sup>2</sup> Whether a given test is significantly less biased than the *Detroit Beginning* or *CTMM Total* is presented in this report. The reader interested in knowing whether other observed differences between means are significant will find such information in West's unpublished M.Ed. thesis.

Where N is not too small (considerably less than 30) those tests which show minimum cultural bias also show, with few exceptions, significant and substantial correlations with academic achievement. Hence, Hypothesis III is supported. Although culture-reduced tests have a minimum dependence of past specific learning, it would appear that these tests do sample from abilities required for academic success.

Hypothesis IV also was supported by the study. Culture-reduced tests by and large have a moderate and significant correlation with other measures of intelligence.

Considering in more detail relationships to the four criteria for evaluation, Table I shows that the SCRIT, Lorge-Thorndike II, Lorge-Thorndike III and Progressive Matrices each satisfy all four criteria at the Grade I level. Differences in extent of cultural bias among these tests are small. However, the SCRIT probably shows greatest promise as a cross-cultural test at this level; it has least bias for both samples.

Table II shows that each of the following tests meet all four criteria at the Grade II and III level: CTMM-Spatial, CTMM-Non-language, Lorge-Thorndike II, Lorge-Thorndike III, SCRIT, Progressive Matrices. The CTMM Non-Language test probably shows greatest promise as a cross-culture test at this level, followed by Lorge-Thorndike II.

At the Grade V and VI level, CTMM Total scores were not available for the white urban comparison group. As indicated in Table III, the Progressive Matrices, the Cattell Test, and SCRIT show least cultural bias at this level.

It will be noted from Table IV that five tests meet all four criteria at the Grade VII and VIII level, and are therefore promising as cross-cultural tests of potential at this level. They are: Lorge-Thorndike I, Lorge-Thorndike III, Lorge-Thorndike Total, Progressive Matrices, Cattell. The Lorge-Thorndike Total as well as having a relatively high mean T-score, correlates highly with achievement and has the highest mean correlation with other tests of the battery. Progressive Matrices ranks after the two Lorge-Thorndike sub-tests.

### Summary, Conclusions, and Implications

The purpose of the study reported here was to investigate a selection of culture-reduced intelligence tests, and to identify those which for two samples of Métis and Indian children, at four grade levels, show a minimum of cultural bias.

A battery of tests, selected by MacArthur for a factor analytic study, was administered to a sample of 126 Métis children attending



TABLE I  
ANALYSIS OF TESTS AT GRADE I LEVEL

Test	Mean T-Score based on Calgary		Standard Deviation		**		*	Correlation with Achievement		Mean Correlation with other tests of battery	
	Faust	Ft. Simp.	Faust	Ft. Simp.	Faust	Ft. Simp.		Faust	Ft. Simp.	Faust	Ft. Simp.
Prog. Matrices	41.4	40.5	6.9	20.9	Yes	Yes	R	.51	.55	.35	.46
S.C.R.I.T.	48.5	41.4	5.9	9.8	Yes	Yes	R	.53	.....	.....	.45
Lorge-Th. III	43.4	39.5	9.8	9.0	Yes	Yes	R	.47	.49	.35	.....
Lorge-Th. II	44.5	38.9	8.2	10.3	Yes	Yes	R	.63	.....	.38	.....
Lorge-Th. I	37.5	32.4	8.4	7.0	No	No	C	.35	.47	.....	.....
Lorge-Th. Total	39.4	34.5	7.7	6.0	Yes	Yes	C	.62	.....	.36	.49
Detroit Beginning											
First Grade	35.0	27.4	4.2	7.6	.....	.....	C	.....	.78	.....	.52

\*Classification = C—conventional; R—culture- reduced.  
\*\*Show significantly less bias than Detroit Beginning, at .01 level.  
Correlations not significant at the .05 level are omitted from the table.

TABLE II  
ANALYSIS OF TESTS AT GRADE II AND III LEVEL

Test	Mean T-Score based on Calgary		Standard Deviation		*	**		Correlation with Achievement		Mean Correlation with other tests of battery	
	Faust	Ft. Simp.	Faust	Ft. Simp.		Faust	Ft. Simp.	Faust	Ft. Simp.	Faust	Ft. Simp.
Prog. Matrices	35.0	36.3	11.1	12.2		Yes	Yes	.72	.30	.58	.34
S.C.R.I.T.	38.4	39.8	8.7	11.1	R	Yes	Yes	.64	.47	.52	.34
Lorge-Th. III	45.2	37.2	10.1	14.1	R	Yes	Yes	.30	.63	.48	.40
Lorge-Th. II	41.3	39.8	8.3	9.8	R	Yes	Yes	.55	.58	.32	.40
Lorge-Th. I	29.4	20.3	7.2	7.4	C	No	No	.54	.64	.52	.40
Lorge-Th. Total	33.0	24.2	9.6	10.3	C	No	No	.65	.30	.41	.34
CTMM-Spatial	47.6	52.9	11.5	10.2	R	Yes	Yes	.43	.41	.50	.35
-N: Lang.	40.8	41.8	12.8	10.2		Yes	Yes	.60	.49	.38	.42
-Logic	35.0	30.0	8.5	6.2	C	Yes	Yes	.60	.35	.55	.38
-Numer.	31.9	26.0	11.0	10.3	C	Yes	Yes	.69	.56	.55	.35
-Total	29.0	21.3	11.5	8.0	C	No	.....	.77	.49	.51	.29
-Verbal	28.0	17.4	8.7	7.4	C	No	No	.74	.56	.50	
-Lang.	25.5	16.5	8.5	5.6	C	No	No	.77	.56		

\*Classification = C—conventional; R—culture-reduced.  
\*\*Show significantly less bias than CTMM-Total, at .01 level.  
Correlations not significant at the .05 level are omitted from the table.



TABLE III  
ANALYSIS OF TESTS AT GRADE V AND VI LEVEL

Test	Mean T-Score based on Calgary		Standard Deviation	*	Correlation with Achievement	Mean Correlation with other tests of battery
	Faust	Ft. Simp.			Faust Ft. Simp.	Faust Ft. Simp.
Prog. Matrices						
Cattell	37.1	39.9	9.7		.62	.71
S.C.R.I.T.	38.4	38.5	10.5	R	.45	.62
Lorge-Th. III	36.0	39.2	8.9	R	.51	.65
Lorge-Th. II	34.1	29.2	15.7	R	.52	.51
Lorge-Th. I	33.6	29.4	11.1	R	.64	.53
Lorge-Th. Total	36.5	34.2	8.5	R	.53	.57
	30.8	26.8	11.0	R	.73	.72
						.53
						.56
						.43
						.62
						.63
						.50
						.61

\*Classification = C—conventional; R—culture-reduced.  
Correlations not significant at the .05 level are omitted from the Table.

TABLE IV  
ANALYSIS OF TESTS AT GRADE VII AND VIII LEVEL

Test	Mean T-Score based on Edmonton		Standard Deviation		**		*	Correlation with Achievement		Mean Correlation with other tests of battery	
	Faust	Ft. Simp.	Faust	Ft. Simp.	Faust	Ft. Simp.		Faust	Ft. Simp.	Faust	Ft. Simp.
Prog. Matrices	40.7	45.2	6.0	7.3	No	Yes	R	.58	.43	.56	.34
Cattell	41.4	40.8	11.0	10.9	No	Yes	R	.56	.38	.47	..
Lorge-Th. III	42.7	46.4	9.9	8.5	Yes	Yes	R	.60	.71	.59	.41
Lorge-Th. II	40.7	44.8	6.8	8.7	No	Yes	R	..	.43	..	..
Lorge-Th. I	46.5	45.0	10.3	12.1	Yes	Yes	R	.48	.58	.45	.39
Lorge-Th. Total	41.9	44.7	8.8	9.2	Yes	Yes	R	.63	.75	.60	.48
CTMM-Spatial	46.0	44.9	7.5	10.1	Yes	Yes	R	..	..	..	..
-N. Lang.	43.4	41.0	7.8	8.4	Yes	Yes	R	.47	..	.48	..
-Logical	37.7	34.8	8.7	9.1	No	No	C	.71	.57	.53	.42
-Numer.	37.1	36.3	9.9	8.9	No	No	C	.56	.59	.48	.40
-Total	36.0	31.4	10.7	10.3	..	..	C	.80	.77	.63	.40
-Verbal	35.9	32.8	12.9	12.0	No	No	C	.74	.71	.45	..
-Lang.	32.8	29.6	12.6	13.1	No	No	C	.80	.75	.47	..

\*Classification = C—conventional; R—culture-reduced.  
\*\*Shows significantly less bias than CTMM-Total, at .01 level.  
Correlations not significant at the .05 level are omitted from the table.



the Faust school. As a replication, an identical battery of tests was administered to a sample of 155 Indian and Métis children attending the Fort Simpson school.

The relative extent of cultural bias in the tests for each sample and at each grade level was determined by comparing the means of derived scores on the tests. All such derived scores were based on a scale of Edmonton or Calgary T-scores in order to assure comparability of performance on the various tests and to permit a comparison of performance between cultures.

The major findings for the study were as follows:

1. Some tests show significantly less cultural bias than do others.
2. Tests hypothesized as culture-reduced show significantly less cultural bias than the more conventional tests.
3. The culture-reduced tests show substantial correlations with academic achievement and therefore sample from the abilities required for academic success.
4. The culture-reduced tests show substantial correlations with other more conventional measures of intelligence and therefore possess concurrent validity.
5. A test which appears to show little cultural bias at one level may show considerably greater bias at another level.

The tests of the experimental batteries used in the investigation were evaluated against four criteria. Those tests which show greatest promise at each level for cross-cultural assessment of intellectual ability have been identified.

The Progressive Matrices is recommended at all levels. The SCRIT also meets the four criteria at the levels for which comparable white urban data were available. Other culture-reduced tests meet the criteria at certain specific grade levels.

#### REFERENCES

1. Eells, E. W. *Intelligence and Cultural Differences*. Chicago: University of Chicago Press, 1951.
2. Elley, W. B. "A Comparative Analysis of Socio-Economic Bias in Selected Intelligence Tests." Unpublished Ph.D. Thesis, University of Alberta, 1961.
3. Garrett, N. E. *Statistics in Psychology and Education*. New York: Longmans, Green and Company, 1950.
4. Hebb, D. O. *A Textbook of Psychology*. Philadelphia: Saunders, 1958.
5. Hebb, D. O. *The Organization of Behavior*. New York: Wiley & Sons, 1949.
6. MacArthur, R. S. "Assessing General Intellectual Ability for Adaptive Teaching Decisions." (Paper presented to Annual Meeting of the Canadian Psychological Association, June, 1961).
7. MacArthur, R. S. "The Coloured Progressive Matrices as a Measure of General Intellectual Ability for Edmonton Grade III Boys." *Alta. J. of Ed. Res.*, 6:2, 67-75, 1960.
8. Vernon, P. E. "The Classification of Abilities," *Educational Research* 3:2, 184-193, 1960.
9. West, L. W. "Assessing Intellectual Ability with a Minimum of Cultural Bias for Two Samples of Métis and Indian Children," Unpublished M.Ed. Thesis, University of Alberta, 1962.

# LEARNING EFFECTIVENESS UNDER THE TRIMESTER SYSTEM

D. A. GIRARD AND F. ENNS

## I. Introduction

One of the concerns of a school system is to improve the effectiveness of learning and, at the same time, to maintain or increase the efficiency of using resources to achieve such improvement. Various means of doing so, such as organizational changes, curriculum modifications or utilization of instructional aids, have been employed. Most innovations in organization, programs or techniques, however, have been introduced within the framework of the conventional school year of ten months.

A further innovation which might contribute to the more effective utilization of resources is the divided school year. Under such a plan the conventional year is divided into two or more terms of equal length. During each term a student completes a limited number of courses. Daily instruction time is adjusted so that although the term is shorter, total time spent is equal to that spent under conventional organization.

The Composite High School in Red Deer, Alberta, has operated since 1949 on such a basis. The year is divided into three terms, or trimesters, of equal length. The terms are independent of each other; instruction in a given course is completed in a single term and a final examination is given at the end of the term. In the next term a student may proceed to a sequent course in the same subject, or to other courses at the same grade level.

The system was originally introduced to accommodate special students who were unable to attend school for the whole year, or who desired to clear a limited number of deficiencies so that they could proceed with further study. The school has functioned successfully over the years but there has been little attempt to make a formal assessment of effectiveness of learning under this form of organization, retention of learning, or improvement in the utilization of staff and physical facilities.

## II. The Problem

Th study<sup>1</sup> being reported here investigated a particular phase of student achievement in the Red Deer School. Specifically, achivement of the Red Deer students on the Grade XII depart-

---

<sup>1</sup> D. A. Girard, "Learning Effectiveness Under the Trimester System of School-Year Organization at the Lindsay Thurber Composite High School, Red Deer, Alberta," Unpublished M.Ed. Thesis, University of Alberta, 1962.



mental examinations in English, social studies and mathematics for the year 1955, 1958 and 1961 was compared with the achievement of students who attended conventionally organized schools and who wrote the same examinations in the same years.

### III. Research Design

The comparison was made by means of a two-way analysis of variance of examination scores of matched groups of Red Deer and control students. The sampling universe was limited to those students who wrote the Grade XII departmental examinations in English, social studies, and mathematics for the years 1955, 1958, and 1961 in the following city school districts: Red Deer, Lethbridge, Drumheller, Medicine Hat and Wetaskiwin. The sampling universe was further limited to include only those students who had progressed from Grade IX to Grade XII in a three-year interval within a given district.

Matching of the groups was done by means of frequency distribution control. The primary trait used in matching was the relevant Grade IX achievement score. The groups were also matched on the basis of Grade IX ability scores, and sex. The effect of socio-cultural factors was controlled to the extent that the Red Deer and the control school communities were similar in these respects. The total population consisted of 274 matched pairs.

The raw mean achievement scores for Red Deer and control students are presented in Table I. The row means indicate the over-all average achievement of both groups in each of English, social studies and mathematics for the years 1955, 1958 and 1961. The column means indicate the average achievement, by years, in English, social studies and mathematics for each of the two groups.

TABLE I  
RAW MEAN ACHIEVEMENT SCORES IN GRADE XII  
DEPARTMENTAL EXAMINATIONS

Year	Subject	Red Deer Group	Control Group	Over-All Average
1955	English .....	55.3	55.0	55.2
	Social Studies .....	60.6	56.3	58.4
	Mathematics .....	50.6	47.3	49.0
1958	English .....	62.6	60.3	61.5
	Social Studies .....	60.2	59.8	60.0
	Mathematics .....	53.2	45.0	49.1
1961	English .....	59.6	58.5	59.1
	Social Studies .....	60.0	59.8	59.9
	Mathematics .....	64.4	51.3	57.9

IV. Analysis of Data

This study was concerned with differences between column means. The row means could be used to compare differences among academic subjects but such a comparison had no relevance to this study. Because there was a possibility that the variance associated with row means would mask the variance associated with column means, achievement scores were standardized to a mean of 50 and a standard deviation of 10 for any matched group. The transformed mean achievement scores are presented in Table II.

TABLE II  
MEAN ACHIEVEMENT SCORES ADJUSTED TO ROW  
MEANS OF 50 AND ROW STANDARD DEVIATIONS OF 10

Year	Subject	Red Deer Group	Control Group	Over-All Average*
1955	English .....	50.1	49.8	49.9
	Social Studies .....	50.1	48.7	49.8
	Mathematics .....	51.5	48.3	49.9
1958	English .....	50.8	49.0	49.9
	Social Studies .....	50.1	49.9	50.0
	Mathematics .....	51.5	47.2	49.3
1961	English .....	50.6	49.7	50.2
	Social Studies .....	50.0	49.9	50.0
	Mathematics .....	54.4	45.5	50.0

\*Error due to rounding appears in third significant figure of row means.

IV. Analysis of Data

The use of standardized scores made it possible to combine the achievement in all subjects for a given year for each of the Red Deer and control groups. The combined average achievement for these two groups is presented in Table III.

TABLE III  
COMBINED ADJUSTED MEAN ACHIEVEMENT IN  
ENGLISH, SOCIAL STUDIES AND MATHEMATICS IN 1955,  
1958 AND 1961 FOR RED DEER AND CONTROL STUDENTS

Year	Red Deer	Control
1955 .....	50.7	49.1
1958 .....	50.7	48.9
1961 .....	50.9	49.3

The data were analyzed by means of a two-way analysis of variance (Schools by Years), using an approximation for unequal frequencies described by Walker and Lev.<sup>2</sup> The summary of the

2 H. M. Walker and J. Lev. *Statistical Inference*. New York: Henry Holt and Co. 1953.



analysis is presented in Table IV which indicates that the over-all differences in achievement between the Red Deer and control students are not significant at the 0.5 level.

TABLE IV  
SUMMARY OF ANALYSIS OF VARIANCE OF  
ACHIEVEMENT IN ENGLISH, SOCIAL STUDIES AND  
MATHEMATICS OF RED DEER AND CONTROL  
STUDENTS FOR THE YEARS 1955, 1958 AND 1961

Source of Variation	Sum of Squares	Degrees Freedom	Mean Square	F-Ratio	F 05
Schools .....	4.16	1	4.16	3.78	3.86
Years .....	0.09	2			
Interaction .....	0.02	2			
Error .....		582	1.10		

The absence of a significant interaction in the analysis of variance further indicates that the relative achievement of Red Deer and control students remained consistent for the years 1955, 1958 and 1961.

As a corollary to this study, the achievement of Red Deer and control students was also compared at different levels. The students to be compared in English and social studies<sup>3</sup> were subdivided into three groups on the basis of Grade IX achievement; those students ranking in the bottom third were classified as a low achievement group, those in the middle third as an average achievement group, those in the top third as a high achievement group.

The achievement between Red Deer and control students within each of these sub-groups was compared by means of a two-way analysis of variance (Schools by Subjects) at each of the three levels of achievement. In no case was there any indication of significant difference in achievement between Red Deer and control students. Furthermore, in no case was there any indication of a significant interaction between schools and subjects.

Conclusions and Implications

The specific research hypotheses tested in this study were designed to assess the achievement of the Red Deer groups in English, social studies and mathematics for the years 1955, 1958 and 1961. Analysis of the data indicated that none of the differences in achievement between Red Deer and control students was statistically significant. The achievement of Red Deer students was at least as good as that of students in conventionally organized schools.

<sup>3</sup> Mathematics was omitted from this comparison because the number of students involved was too few.

However, the study was confined to those students who had progressed normally through the high school grades. It did not assess the achievement of those students who might have found the trimester system attractive for a variety of special reasons. Therefore, the findings are relevant to a limited aspect of the trimester organization only.

Notwithstanding this limitation, the study performed a useful function. Any innovation must meet the needs of the normal student as well as those of the special student. That this can be done in the divided school year was demonstrated. Students who had progressed through the high school program in three years achieved as well under the trimester organization as in the conventional school.

For complete appraisal of the divided school year, other aspects beyond the scope of the present study should be investigated. The following are suggested for further study:

1. *The Special Student.* The apparent flexibility of the trimester system suggests that special needs of students can be met. The way in which this is achieved for the academically gifted student, the student requiring remedial treatment or the student of mature years requires study and appraisal.

2. *Learning Effectiveness.* While the present study examined achievement on Grade XII examinations, it would also be valuable to compare retention of knowledge. The shorter, more intensive periods for studying any one course suggest that there is less time for the student to assimilate or integrate his learning and therefore that the rate of forgetting will be higher than among students in conventional schools. The present study suggested that this may not be so. Red Deer students performed as well on examinations as their counterparts after three years exposure to the trimester system. To the extent that success on the Grade XII examination depends on retention of learning from Grades X and XI Red Deer students' retention of knowledge was as good as that of students from conventional schools. This conclusion requires further testing, however.

The divided school year is a significant experimental innovation in secondary education. This is one way of varying the "input mix" to try to increase desirable output. While it is not suggested that all secondary schools should adopt the divided school year, such a plan may help a particular school to discharge its obligations more adequately.



## REFERENCES

- Edwards, A. L. *Experimental Design in Psychological Research*. New York: Rinehart and Company, Inc., 1960.
- Garrett, H. E. *Statistics in Psychology and Education*. Toronto: Longmans, Green and Company, 1958.
- Girard, D. H. "Learning Effectiveness Under the Trimester System of School-Year Organization at the Lindsay Thurber Composite High School, Red Deer, Alberta." Unpublished M.Ed. Thesis, University of Alberta, 1962.
- Guilford, J. P. *Fundamental Statistics in Psychology and Education*. Toronto: McCraw-Hill Book Company, 1956.
- Johnson, P. O., and R. W. B. Jackson. *Modern Statistical Methods*. Chicago: Rand McNally and Company, 1959.
- Lindquist, E. F. *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton Mifflin Company, 1953.
- Selltiz, Claire, and others. *Research Methods in Social Relations*. New York: Henry Holt and Company, Inc., 1960.
- Walker, H. M., and J. Lev. *Statistical Inference*. New York: Henry Holt and Company, 1953.

# AN EXPERIMENT IN TEXTBOOK SELECTION FOR GRADE TEN LANGUAGE

JAMES B. BELL

*Superintendent of Schools  
Alberta Department of Education*

## Problem

### *The Purpose of the Study*

This study in textbook selection was conducted at the request of the Subcommittee on High School English, Curriculum Branch, Department of Education for the purpose of obtaining, recording and reporting teacher opinion on the effectiveness of certain pre-selected textbooks in the grade ten language program for Alberta schools. *Creative Composition* (4) was to be judged by teachers in terms of its effectiveness as a language text that might serve as an alternate to, or replacement for, the currently used language text, *English for Today*. (3) *Basic Spelling for High School Students* (1) and *English Composition Book II* (5) were to be evaluated in terms of their effectiveness as possible parts of, or supplements to, the grade ten language program.

### *Importance of the Study*

In past years the Curriculum Branch of the Department of Education relied largely on the members of the subcommittee to determine whether or not a text was suitable for use in a certain grade. This study, then, because of its reliance on the opinion of classroom teachers, represents a departure from the usual procedure followed by the Subcommittee on Senior High School English in selecting textbooks to be authorized for use in Alberta schools.

As the purpose of this study was to obtain teacher opinion on the whole as well as the individual parts of certain texts, it was believed that the questionnaire method to be employed should include a combination of objective and open-end questions that would produce the accuracy required and still allow for free expression of opinion.

In addition to the development of a questionnaire method for the pilot project, it was necessary to plan a procedure that would ensure the most complete and accurate expression of teacher opinion.

The importance of the study is based on the following factors:

- (1) the trend in Alberta toward greater teacher participation in textbook selection.
- (2) the desirability of devising a particular questionnaire method for obtaining teacher opinion on the effectiveness of textbooks in a classroom situation.



- (3) the need for the development of a procedure that might serve as a guide in future projects of this nature.

### *Limitations of the Study*

This study was limited to the consideration of certain pre-selected textbooks that were judged by the Subcommittee on Senior High School English as most likely to contribute to the development of the language program in grade ten.

The number of teachers participating in the study was limited to twenty, these teachers being chosen on the basis of qualifications and competence by the Subcommittee. The participating teachers were not asked to assist in the selection of textbooks for the survey, or to do analytical study, but only to make an evaluation, the scope of which was determined by the questionnaire.

The related literature dealt with principles of textbook selection and with the ten areas of an effective language program most frequently emphasized by writers in the field of language teaching. The attention of participating teachers was not called to the review of related literature.

The answers to the questionnaire were assembled and interpreted by the writer. Teacher opinion was reported to the subcommittee members who made the final decision as to whether the texts should be authorized as replacements, alternates, suggested references, or whether they should be used at all.

### *Definition of Terms*

The Subcommittee on High School English is a group of professional educators (teachers, school inspectors and university professors) chosen by the Department of Education to evaluate existing courses, to develop new courses, and to recommend textbooks.

The Faculty Committee is made up of the university personnel who serve as members of the Subcommittee on High School English.

Multiple Authorization of textbooks is the authorization by the Department of Education of two or more textbooks for the use of teachers and pupils in any course. School systems may use one or more of these texts.

## **Procedure**

### *Plan of Study*

Two supervisory visits were conducted during this study. The first visit was made during school opening week to explain the purpose of the study, to discuss the use of the textbooks, and to allow participating teachers to ask questions. The second supervisory visit took the form of a structured interview. The questions were designed to obtain teacher opinion on the effectiveness of the text-

books and to gain information and impressions which might assist in structuring a questionnaire. Reports, both oral and written, were made following each supervisory visit to the Subcommittee on High School English.

In constructing the questionnaire, the investigator first surveyed the opinions of authorities on textbook selection. Most of these opinions were related to the two steps suggested by Waterman (6). These two steps are:

1. The formulation of a set of criteria on standards by which textbooks under consideration may be judged.
2. The conduct of comparative studies, objective in nature as far as possible, to determining the relative merits of the several books on each item of the criteria.

By focussing attention on a number of recent representative articles, periodicals and books, it was possible to derive a list of ten areas considered by most authorities to be important in any language program. These important areas provided criteria of the relative merits of the two texts, *Creative Composition* and *English For Today*.

The criteria are listed below:

1. Provision for Individual Differences
2. Motivating Students
3. Preparation For Students' Writing
4. Provision For Creative Writing
5. Provision For Expository Writing
6. Developing Clear Thinking
7. Increasing Students' Skill In Organizing Thought
8. Increasing Students' Skill in Writing Accurately
9. Developing Vocabulary
10. Using Grammar

Open-end questions were employed to obtain information of special interest to the subcommittee and to allow teachers to express freely their opinion on the text or any part of a text.

### Conclusions

1. The material in *Creative Composition* was better suited to the the students of above-average and average ability than to students of below-average ability.
2. *Creative Composition* was the kind of language text that would help the inexperienced teacher in dealing with individual differences in the classroom.



3. The teachers believed that various motivational devices used in *Creative Composition* were successful in creating within the student a desire to improve his writing. Of these devices the most successful seemed to be the use of models, and the least successful was the writing-partner idea.
4. *Creative Composition* was effective in assisting the teacher to prepare students for writing.
5. The emphasis on creative writing in *Creative Composition* met with the approval of English teachers.
6. *Creative Composition* did not develop expository writing to the extent that some teachers believe it should be developed.
7. Clear thinking, vocabulary building, accurate writing and organization of thought were satisfactorily developed in *Creative Composition*.
8. The grammar in *Creative Composition*, although of a traditional nature, was functional and well developed. However, it was too elementary for the better students who required more mature supplementary material.
9. Ninety-five percent of the participating teachers believed that *Creative Composition* was "as effective", if not "more effective" than *English For Today*.
10. The omission from the language program of a unit on mass media of communication was not considered a significant loss.
11. *Basic Spelling For High School Students*, or a somewhat similar spelling textbook, would be considered a valuable addition to the grade ten language program.
12. Although the need for writing models, reading comprehension, vocabulary and grammar exercises was frequently stated in answer to open-end questions, it was not believed that the text *English Composition Book II* satisfied these needs in a manner that would be generally accepted by teachers and students.
13. The use of a set of standards based on important areas of a language program, plus a series of open-end questions produced a clear complete impression of teacher opinion concerning the effectiveness of *Creative Composition*.

### Recommendations

1. The Subcommittee on High School English might give further consideration to the materials necessary in building a suitable language program for the below-average students, to the testing of these materials in certain classrooms, and to the recommending of useful materials to classroom teachers.

2. The subcommittee should give further consideration to whether or not one spelling book should be authorized or a number of possible books should be suggested for use in grade ten classes in Alberta schools.
3. In view of the comments made by a number of teachers and the stress placed by authorities on expository writing, the Subcommittee might give further consideration to the balance between expository and creative writing in a grade ten language program.
4. At present the Curriculum Guide states that *Guide to Modern English* (2) is authorized for use in grades eleven and twelve, but may be used in grade ten at the discretion of the teachers. Since this text is considered a valuable supplement and guide for the average and above-average students in grade ten, the Department of Education might give consideration to its authorization for use in grade ten.
5. In view of the emphasis placed by teachers on the need for reading comprehension exercises, the Subcommittee might give consideration to recommending for authorization certain supplementary material that would assist teachers in dealing with those students who have comprehension difficulties.
6. Future experiments of this nature should be more tightly controlled through
  - (a) A careful selection of participating teachers according to qualifications, experience, competence and teaching situation.
  - (b) A carefully planned use of certain textbook combinations.
7. In view of the trend toward increased teacher participation in textbook evaluation and selection, the Faculty of Education, University of Alberta should consider giving teachers-in-training instruction in the research, and the methods and principles of textbook selection.
8. The Department of Education should continue to encourage teacher participation in textbook selection.

## REFERENCES

1. Bowden, N. J. *Basic Spelling For High School Students*. Toronto: The Macmillan Company of Canada Limited, 1960.
2. Corbin, R. K., Perrin, P. G., Buxton, E. W. *Guide to Modern English*. Toronto: W. J. Gage Limited, 1955.
3. Gray, M., Hatch, C. W., Meade M., Waddell, W. *English For Today*. Toronto: Longmans, Green and Company, 1954.
4. McMaster, R. J., McMaster, W. C. *Creative Composition*. Toronto: Longmans, Green and Company, 1957.
5. Scott, A. F. *English Composition Book II*. Toronto: The Macmillan Company of Canada Limited, 1960.
6. Waterman, I. R. "When You Choose a Textbook". *The Phi Delta Kappan*, XXXIII, 1951-52, pp. 267.



# PROBLEMS INHERENT IN ABILITY GROUPING

ARTHUR A. ATTWELL

*Los Angeles State College*

AND

THOMAS E. LINTON

*University of Alberta, Calgary*

Ability grouping in the senior high school setting is a not uncommon practice. Generally the literature supports the idea of using different classrooms for the separation of ability groups at this level.<sup>1, 4, 5, 11</sup> However, when one considers the value of similar grouping at the junior high and elementary levels, certain important differences appear. The research in these areas tends to lack the clarity of implication which is more often found in the articles on senior school grouping.

This paper will examine some of the less commonly measured variables in homogeneous grouping at the intermediate school level.\* Previous studies have considered those factors which are more conducive to standard measurement techniques. The trend has been toward the assessment of school achievement as the result of ability grouping, rather than evaluating the personality variables which are also an important part of the grouping process.

Before looking at these variables it is appropriate to consider briefly the advantages which are often stated for the ability grouping program.

1. The extremes of academic ability are removed from the "average" classroom.<sup>10</sup> This removal permits the teachers to concentrate more fully on the relatively similar level of ability which remains in the room.
2. Curriculum adjustments in an overall planning sense are made more feasible by the differentiated classes.<sup>7, 11</sup>
3. The more able in terms of "giftedness" or high achievement are not restricted or forced to adjust to the lower capacity levels of other class members.<sup>5, 8</sup>
4. The lower ability group is not subject to personal and group comparisons as the result of a more openly competitive situation.<sup>10, 11</sup> Hence the "dull-normal" student does not directly lose status when his achievements are viewed against those of other individuals of similar abilities.<sup>10</sup>
5. The attitudes of the children towards school improve at both ends of the ability continuum.<sup>9</sup> It is believed that this improvement in attitude leads to an increase in academic motivation.<sup>10</sup>

There are other factors in addition to the stated advantages, which may directly affect the outcome of the grouping process. The authors feel that these secondary factors play an important part in determining the success or failure of ability grouping in a particular school.

What are these factors and what role do they play in advancing

---

\*Grades six, seven and eight.

or limiting the use of the homogeneous teaching methods? These points will be presented as questions, and commented on at this point.

(1) *Is Homogeneous Grouping Actually Possible?*

*Comment:* The literature on grouping is not all favorable. Some of these same questions were asked in the pioneer studies undertaken in the 1930's. Burr has described the extreme difficulty in even attempting to group children by ability.

Individual pupils are not themselves homogeneous in physical or mental traits, or in achievement in school subjects. An average achievement score represents a mean of specific achievements which may vary so widely that groups relatively homogeneous in terms of average educational age are relatively heterogeneous with respect to achievement in any particular subject.<sup>2</sup>

Not only are the achievement levels different in different subject areas, but the actual range of abilities in the homogeneous class may be very large. The group tests do not adequately measure the higher I.Q. ranges. They do not sufficiently differentiate between an individual's range of ability in an area, and the modal number of responses requisite for placement in a "high" ability group. Hence the ability group may represent the high achievers, and the socially more adaptive, rather than the actually more capable academically. This may be the product of social class factors operating at the school level. The group test and the "homogeneous" teaching approach may promote the interests of the upwardly mobile parent, and permit the administration the indulgence of "meeting parental needs" without actually determining the distribution of ability within the school setting. In this sense ability grouping may represent social class factors, rather than a real assessment of student capacities.

(2) *Does Segregation By Ability Produce A Positive Feeling For The Individual Students, Or Does It Help To Promote Attitudes Which Are Inimical To A "Balanced" Social Development?*

*Comment:* Keliher has questioned the use of I.Q. and achievement tests as an acceptable basis for the grouping procedures.

Grouping concerns and affects the whole personality, as does segregation into bright, average and dull groups. Grouping may cause serious attitude hazards which might be investigated further,—so far as one can study the heretofore unmeasured consequences of homogeneous grouping—segregation has dangers for mental health, and therefore, in the light of these dangers,—is not desirable.<sup>3</sup>

Homogeneous grouping on the basis of achievement tests, along with segregated schools for the different levels of academic achievement have long been a part of the European educational approach. The assumptions underlying this pattern of student selection are based in the aristocratic tradition which has characterized European life for centuries. If ability grouping pro-



motes social class differences, then the grouping process needs to be questioned when used at the earlier levels. For the Canadian schools do not intentionally foster classroom arrangements which deny individual worth.

- (3) *Does The Child Become Aware Of His Placement In The "Superior" Or Below Average Group? Does The High Ability Group Tend To Develop "Superiority" Feelings To A Greater Degree Than Would Be The Case In The Heterogeneous Situation?*

*Comment:* Just as the child becomes aware of social class differences in the early elementary grades, so he recognizes the meaning of ability grouping in the early years. How he reacts to this knowledge in behavioral terms is difficult to say. It is not difficult to estimate the child's feeling of personal worth after he becomes aware of the difference between himself and his classmates. Even when the academic advantages are granted in the homogeneous placement one wonders whether the personal adjustments for the pupils are worth the problems they tend to create. The point is not merely one of academic competence, but also of human competence which is more resistant to measurement. Since personality growth is felt to be of major import in the elementary years, it is at this age that these questions should be viewed as central determinants in the establishment of the grouping program.

In Canada "superiority" is viewed as a many sided quality. One is not superior merely on the basis of academic achievement. While the more academic professors may grind their teeth over this social tradition it is nevertheless is a characteristic of Canadian society. This social tradition in Canadian life has not fostered an élitist social group with similar values and attitudes. In these terms the heterogeneous classroom, in the early school years, promotes greater social understanding and acceptance of the differences between men, rather than increasing the basis for conflict by encouraging an attitude of superiority and inferiority among school children.

It is the impression of the authors that children in the lower groups are well aware of their placement, and tend to feel some resentment toward the higher ability groups.

- (4) *Does The Lower Ability Group Have A Sufficient Leadership Potential In Order To Establish The Necessary Norms For Adequate Academic And Social Behavior In The Classroom Situation?*

*Comment:* It seems that when the lower ability groups are isolated from the other ability levels, the leadership capacity is

also removed. This absence of leadership tends to create a social and academic vacuum which the teacher is personally unable to fill. The result is that social and behavioral "norms" are very difficult to establish, and control is harder to maintain. The lack of leadership within the room causes a lack of social cohesiveness in the group, inhibiting "self-discipline" in the class, and necessitating a teacher-imposed kind of control.

(5) *Is There A Tendency For The School's Discipline Problems To Be Placed In The Lower Ability Groups?*

*Comment:* It is a not uncommon practice in the schools for certain non-academic areas of the curriculum to become depositories for the least capable students. These are favored and non-favored parts of the school program. These areas do not necessarily correspond to the actual ability of a particular child. A given student may be very "bright," but his behavior in the academic classroom is disruptive and unmanageable. He is subsequently placed with the low achievers. The result is that the lower ability groups contain some students who are very able, but are placed there as the result of a system which rewards certain patterns of social behavior. This lower placement serves to reinforce the socially unacceptable behavior.

The behavior problem and the low achiever are frequently placed in the lower group on the basis of marks and conduct which are often interrelated. Once the student develops an image of himself as an "untouchable" in a particular school, his reputation tends to run ahead of him, creating a negative "halo" effect among the faculty. This serves to label his capacities, and make it more difficult for him to actually achieve at a higher level.

(6) *Is Higher Achievement A Function Of Ability Placement, Or Would The Individual Do As Well In A Heterogeneous Situation?*

*Comment:* A study by Abramson<sup>1</sup> indicated that ability grouping produced "no significant differences in the grade point averages of students in ability groups and equal students not grouped on an ability basis." His conclusion was that the overall achievement of students, as indicated by a grade point averages and honors, was associated with their level of intelligence, rather than with the particular ability group they attended. A study by Wrightstone reported essentially the same findings.<sup>12</sup>

(7) *Do Parental Attitudes And Pressures Play An Important Part In The Successful Outcome Of The Grouping Program?*

*Comment:* Historically, ability grouping has followed a repeti-



tive pattern. It has been socially favored, even demanded by parental pressure groups. Then a period of disenchantment follows when parents are unhappy with the results, for they soon realize that their "Johnny" may be in a lower group, and socially marked as an incompetent child. If the child is poor academically then the parents may also feel inadequate and hostility to the school program may follow. The result is that the program is altered to meet the changing demands of the parents, and finally the intent of the grouping process is lost in adjusting to community pressures.

However grouping may be more successful when the parents are carefully approached and involved in the ability program. This may mean adult education of the more informal type to make the parents aware of the problems involved in the ability grouping process.

(8) *Are Teachers Negatively Affected, Or "Demoralized" When They Remain As Teachers Of The Lower Ability Groups?*

*Comment:* There does seem to be a negative and "demoralizing" effect upon the teachers who remain with the lower groups for a prolonged period of time. This is caused by several related factors. A socially negative value is attached to the teaching of the lower groups by the teachers in the upper groups. The social "pecking order" in the schools tends to place the high academic groups as the most worthy, and their teachers as the most able in the school. Hence a teacher who continuously works with the low ability group may be ranked as inferior in this status system. Academically the amount of learning is less in the lower group, and the feeling of satisfaction with pupil accomplishment may also be less. The lower groups are more difficult to control and less amenable to standard teaching procedures. This means that the teacher does more work for the same rewards, and receives less prestige in the school's social ranking system.

Hence, unless the teachers are given the opportunity to teach other types of classes they may become unhappy with their situation, and lessen the programs effectiveness. This demoralized feeling may not be a problem for the specially trained teacher whose skills are more highly valued by the other members of the staff.

*Summary Comment:*

When ability grouping first become popular during the 1920's two of the major pioneer studies suggested that there might be some problems inherent in the process. These problems considered the difficulties of actually grouping children on a homogeneous

basis, and the resulting attitudes of the students placed in these groups. It was also wondered whether consideration was given to the mental health aspects of this kind of grouping, and whether there might be any unmeasured consequences which would be inimical to the best interests of the developing individual. While research on the subject is divided, the popular trend favors ability grouping. Most of the present studies measure academic progress, and seem to be unconcerned with the problems discussed in this paper.

The tendency to accept grouping at the upper junior high level may in part reflect the popularity of this approach in contrast to its lack of acceptance in the elementary school years. This rejection during the early school years has been a traditional feature in a value system which essentially underlies elementary school practice. These values argue that placement is antidemocratic and not in the best interest of the child's personal development. It is felt that personal and emotional growth should take place in a "normative" classroom structure. The assumption is that grouping is a "non-normative" social situation, and that it produces negative effects on the student's personality development. While this assumption has not been established on the basis of empirical research, it has become the "popular" answer for heterogeneous grouping at this level.

In the secondary schools grouping is more commonly found partly because of a different set of values justifying its use. These values argue that "academic" material may be presented more intensively and coherently in the higher ability situation. The emphasis is on academic achievement at this level and not on personality adjustment. These values are not products of research, but rather are part of a traditional approach which holds that if it's "academically oriented" it must be correct.

Grouping as a process has seldom been advanced on the basis of "research." One reason for this is simply that the empirical studies do not clearly evidence a strong position either way. Individual academic progress in the ability group setting does not appear to be any greater than in a heterogeneous group. In addition there are the social and emotional factors which have been presented in this paper. The authors feel that both the academic aspect of achievement, as well as the personal growth factors must be considered when grouping procedures are undertaken.

#### REFERENCES

1. Abramson, David, "The Effectiveness of Grouping for Students of High Ability," *Educational Research Bulletin*, Oct. 14, 1959, 169-182.
2. Burr, Marvin Y., "A Study of Homogeneous Groupings," *Columbia University Contributions to Education*, No. 457, New York, 1931.



3. Cornell, E. L., "Effects of Ability Grouping Determinable from Published Studies," *35th Year book of the National Society for the Study of Education*, 1936, 298-302.
4. French, John W., "Evidence From School Records on the Effectiveness of Ability Grouping," *Journal of Educational Research*, Nov. 1960, 82-91.
5. Hansen, Carl F., "Ability Grouping in the High Schools," *Atlantic Monthly*, Nov., 1960, 123-127.
6. Keliher, Alice V., *A Critical Study of Homogeneous Grouping*, Columbia University Contributions to Education, No. 452, New York, 1931.
7. King, James, "To Group or Not to Group," *Ohio Schools*, Dec., 1960, 18-19.
8. Kinkaid, Donald, and Epley, Thelma, "Cluster Grouping," *Education*, Nov., 1960, 136-139.
9. Loomis, Mary Jane, "The Right Child in the Right Classroom," *NEA Journal*, Sept., 1959, 17-18.
10. Savard, William, "An Evaluation of the Second Year of an Ability Grouping Program," *California Journal of Educational Research*, March, 1961, 62-66.
11. Watson, Robert, "Let's End the Grouping Debate," *The American School Board Journal*, July, 1961, 44-47.
12. Wrightstone, J. Wayne. "Class Organization for Instruction," *NEA Journal*, April, 1957, 245-255.

# LEISURE READING IN THE SENIOR HIGH SCHOOLS OF ALBERTA

C. S. H. CAMPELL

In 1959 a survey of leisure reading in junior and senior high schools of Alberta was undertaken by two members of the Leisure Reading Sub-committee of the Department of Education. The Committee on Educational Research, University of Alberta, provided the financial assistance.

A province-wide survey of high school leisure reading was therefore undertaken by means of two questionnaires, one sent to all the teachers of high school leisure reading in the province, and the other to a proportional sample of approximately 10% of the students in grades seven to twelve.

The responses, approximately 1,100 from teachers and 7,700 from students, were transferred to IBM cards. The data from the junior high schools were analysed by Mr. R. R. Fisk, and, in 1961, accepted by the University of Alberta in partial fulfilment of the M.Ed. degree. The data obtained from the senior high schools, separate but complementary to that of Mr. Fisk, forms the material of this article.

## General Purpose

The purpose of the survey was to show the strengths and the weaknesses of the Alberta high school leisure reading program as they were reported by the teachers and the pupils engaged in that program. The investigator wished to discover as precisely as possible what books senior high school students were reading, whether the Department of Education objectives in leisure reading were being met, and, in general, what attainment was being made in high school reading.

## Hypotheses

The investigation centered around five specific hypotheses:

1. Students' attitudes and behaviours are affected by their leisure reading.
2. The leisure reading program can be implemented in various ways.
3. Various observable factors affect leisure reading accomplishments.
4. Leisure reading accomplishment varies greatly.
5. Changes are necessary in the present leisure reading program in Alberta.



### Areas of Investigation

The students' questionnaire did not elicit data on the students' abilities (I.Q. ratings), their background, or the library facilities of their schools, but it did obtain information of the type of school attended, the age, grade and sex of the respondents. The teachers' questionnaire obtained extensive information on teaching methods, and the teacher's evaluations of student achievement and of the leisure reading program. The sex, the teaching experience, the knowledge of books and library facilities of the teacher respondents were secured also.

### Procedure

The first instrument used contained fourteen questions on such aspects of the students' leisure reading as: student reading interests, factors affecting reading accomplishment, reading habits, gains from, and improvements necessary in the established leisure reading program. The second instrument was composed of thirty-two questions on aspects of leisure reading within the knowledge of teachers. After pilot studies had established the suitability of the two instruments, the student questionnaire was sent a representative sample of about 10% of all high school students in cities, towns, divisions, villages and rural areas; and teacher questionnaire to all teachers of leisure reading in the province.

The findings from each questionnaire were presented separately. From the 3,500 replies of students of the senior high school 1,000 were chosen at random. These responses were tabulated and discussed generally according to the type of school the students attended: Type A—city schools with population of 500 and more; Type B—schools with population from 100 to 499; Type C, population 50 to 99; Type D, small schools; and Type E, Roman Catholic Schools. Additional findings were tabulated and discussed under grades (and ages) and sex.

The responses from all the teachers were tabulated and discussed under three general headings: methodology in leisure reading, evaluation of student achievement, assessment of the leisure reading program. As teachers' experiences seemed to produce more variation in response than differences in sex, knowledge of books, library facilities, etc., this factor was presented at greatest length.

### Findings

#### *From the Students' Questionnaire*

*Findings According to Type of School.* High consistency of responses from the students attending the five types of schools was obtained. A few conflicting opinions were noted, however. As the sample by schools was randomly and proportionately selected from

across the whole province, the following generalizations seems to be valid regarding the leisure reading abilities, habits, requirements of senior high school students.

High school students require (but do not desire) adult guidance in choosing books and types of literature. More can be gained from a guided leisure reading program than is presently realized.

The influence of mass media of communication (television, moving pictures and advertising in magazines) with their emphasis on crude types of realism is making some impression on senior high school students. This influence varies considerably from school to school but it is not in any case great,—far from devastating as some writers would imply.

Library facilities, which vary greatly across the province, seem to be inadequate to supply the wide reading requirements of senior students.

Reading preferences seem to be firmly entrenched across the province.

Though reading appears not to be a habit with Alberta youth, it competes strongly with other kinds of recreation and with numerous distractions for a share of the students' time. The students on the average claim to read seventeen books per year.

Classics seem to have maintained favor with students in Alberta schools, in spite of suggestions in the related literature that they are not always regarded as suitable for some United States students.

*Findings According to Grades and Ages.* Clear evidence was obtained that most students in reading show a definite growth toward maturity; nevertheless many do not form the habit of reading better-quality novels. They prefer sports and mystery stories and lighter magazine reading. But the trends toward maturity, which are itemized below, seem to outweigh the opposing tendencies.

As the students become older, a preference for books featuring adult rather than juvenile heroes is evident.

With increase in age a decreasing interest in stories and books of love, hobbies, careers, personal problems, Western adventure, fantasy, animals and science is evident.

A slight increase in interest in plays is evident, especially from age sixteen on.

Though a decrease in the number of books read per year after age sixteen is noticed, this trend is probably the result of greater demands made on the students by employment and by high school work.

As the students grow older, a decrease in interest in comics, movies and television is noticed.



"Thrills" from reading appears to give way to "curiosity," as the students advance through high school.

A maturer attitude to teacher influence, and to writing book reviews seems to develop as the students grow older.

A desire for "adult" books and for books adults read seems strongly evident among older students.

With increase in age reading appears to become more a hobby, a habit, a fixed and regular recreation.

Powers of evaluation and literary appreciation appear to increase with age.

Less identification with the heroes of novels seems to occur as the students approach maturity. This tendency may be interpreted to mean that the readers regard the novel as a work of art.

As the students get older, a growing interest (or a contemporary interest) in travel seems to develop.

Among the older students not much interest is shown in respite or escape literature.

For the older students reading appears to be done because of its value, not because it is expected of them or because it thrills them.

*Findings According to Sex.* Marked differences were discovered in the male and female attitudes to books and to the leisure reading program. But just as many, or more, close correspondences between the responses of males and females to each question of the questionnaire were noticed. Following is a summary of the main differences and the main similarities according to sex.

#### (a) Differences

Non-fictional books about occupations and about grooming or personal appearance rated high with girls, low with boys. Descriptions of war and conquest, and explanation of invention and scientific discoveries were much preferred by the boys but were uninteresting to girls.

Love stories, tales of personal achievement and career stories were preferred by the girls, but disregarded by the boys. They, in turn, liked sports stories and science fiction, which the girls did not care for.

Boys preferred essays to poetry and plays; girls showed the reverse preference.

Girls, apparently more avid and more mature readers than boys, reported reading to be a pleasure in slightly higher percentages. Boys, in greater percentages than girls, asserted that reading was not a pleasure. It would appear that girls like a passive recreation and that boys prefer activity.

Boys appeared to read for the purpose of gaining information on such topics as their hobbies, or to satisfy curiosity. Girls more often than boys appeared to like to identify themselves in various ways with the characters or situations in the novels they read. It appeared that boys hoped for practical values from their reading; girls were more apt to look for artistic values or for opportunities to identify themselves with people in literature.

It must be noted that the labels "boys" and "girls" do not clearly distinguish two separate categories of readers. Some boys are not typically "boyish" and many girls are not unreservedly "girlish".

### (b) *Similarities*

On almost all of the general aspects of leisure reading,—the methods of choosing books, the difficulties encountered in completing the leisure reading requirements, students reading habits, and their reasons for reading, the gains from the leisure reading program, and the improvements they suggest—boys and girls showed a high degree of agreement. For example, all preferred magazines to poetry and plays. Again, they agreed, within half a book per student per year on the number of books read per year. And a similar percentage of boys and girls stated that they like leisure reading.

### *From the Teachers' Questionnaire*

The survey was based upon the opinions of over three hundred teachers, of both sexes (equally), working in all types of schools. They had varying amounts of experience; half of them had taken University English courses and some had read widely. Some worked with quite elaborate library facilities, others with quite inadequate supplies of books. Their responses to all questions on the questionnaire showed a high degree of consistency.

In methodology the teachers preferred a controlled or directed program to a "free" leisure reading program. They encouraged and urged but did not care to force their students to read. They checked their students' leisure reading accomplishment quite closely, most frequently by means of written book reports, which dealt both with the structure and the literary values of the books. Teachers preferred to use informal advice and casual recommendations rather than more direct or obvious forms of motivation.

Teachers seemed to agree, with some reservations, that student accomplishment in leisure reading was quite good; but they also believed that more could be accomplished. The objectives of the leisure reading program were, according to the majority of teachers, being met by a majority of the students.

Teachers similarly agreed that the program itself was generally



satisfactory. They appeared to know their students' reading interests; they saw definite gains accruing from leisure reading, but they recognized many minor problems still to be solved if a highly effective program is to result. It appears that they recognized many of the factors which will (or should) produce improvement in leisure reading in Alberta high schools. Furthermore, they seemed willing to accept help or suggestions if aid were made available.

*Areas of Agreement between Students and Teachers*

Not only were the students' and the teachers' response consistent within themselves, but, where parallel information was sought from each group, students and teachers, answering quite independently, agreed on many aspects of leisure reading. The teachers, for instance, agreed with the students on the type of fiction and non-fiction that they, the students, favored. Both teachers and students made the strongest plea for more books. In fact, both groups stated that a successful reading program could hardly be conducted unless the number of books available in school and classroom libraries was increased. Teachers believed that students were gaining vicarious experience of life from their leisure reading; the students, in turn, reported that they gained knowledge of other people and other places. The teachers reported that students were improving in reading skills. The students, similarly, reported that, though they did not believe that they were gaining greatly in literary appreciation, they were improving their knowledge and their reading ability.

The difficulties in leisure reading reported by students paralleled to some extent the problems and adverse effects reported by the teachers. For instance, both groups reported that the demands of school work and home duties, of other forms of recreation and the effects of distractions caused students most difficulty in their leisure reading.

Approximately 90% of the students reported that reading was a pleasure to them. About 70% of the teachers reported that over half of the students found reading "definitely pleasurable". Teachers further reported that most of the students "benefit considerably" from the leisure reading program, a claim that was corroborated by the students.

It appeared that students and teachers agreed that the leisure reading program was of benefit. They seemed, further, to agree that the program was largely satisfactory, though this agreement was merely implied in teacher and student responses as a whole. No sweeping claims of benefit or satisfaction were made by either the student or the teacher groups. In fact the claims were moderate and modest. A small percentage of both teachers and students dis-

agreed with the more commonly reported opinions, but teachers and students were in agreement rather than in opposition on all general aspects of leisure reading.

### Conclusions

In terms of the "General Purpose" mentioned above, the following conclusions seem warrantable from the data of the study.

The students of the senior high schools of Alberta are reading quite widely, and are profiting from the time spent in leisure reading. They report that they read an average of seventeen books per student per year. The students recommend hundreds of books from the reading lists and hundreds more not on the reading lists. Teachers, assessing the students' accomplishment, report that the students on the average meet about 75% of the Departmental requirements.

Partly because of these requirements but partly of their own volition, Alberta high school students accept reading as a regular recreation, though the demands of work and amusements compete with it strongly for student attention.

A great variety of books is being read by students. Many adult books, book condensations for adults, paperbacks and junior novels, those listed in *Invitation to Read* and many others found elsewhere, were recommended by students. Many of the books listed in *Invitation to Read* did not appear to be popular among high school students.

The objectives of the leisure reading program as stated in Department of Education bulletins are being achieved by the majority of the students according to the report of the majority of the teachers.

It appears from both teacher and student responses that a good deal of value is accruing from the leisure reading program.

The objectives of the Alberta leisure reading program seem, judging from the tenor of the students' and the teachers' responses, to be reasonable and attainable.

The leisure reading program appears to be improving the mental and affective processes connected with reading.

In terms of the "Hypotheses" mentioned above, the conclusions of this study partially support Hypothesis 1 and strongly support Hypotheses 2, 3, 4 and 5.

Some conclusions arising from Hypothesis 5 (changes necessary in the leisure reading program) and some arising from findings throughout the investigation may be of interest. For instance, the gap which appears to exist between junior and adult novels did



not appear to be adequately bridged either by books from the approved list or by other books discovered by the students. Though the junior novels listed in *Invitation to Read* were frequently read by senior high school students, there were indications which suggest that they are somewhat insipid and stereotyped. The harsher realism of war stories and modern psychological studies of maladjustment seemed to appeal more strongly to students than did the simpler, milder junior novels to be found in school libraries.

Many more books and many more titles were, of course, badly needed in school libraries.

More guidance was required in the selection and study of books for Alberta school libraries. The teachers indicated by their responses that more care was required in the selection of books listed in *Invitation to Read*. Better annotations, and grading of books according to difficulty were also necessary.

Teachers' opinions support the conclusion that further restrictions within the program are unwarranted. The middle ground between a completely "free" program and a completely prescribed program seems best, and the Alberta program lies somewhere in this region.

To increase the quantity and to improve the quality of books read by students, and to increase teachers' enthusiasm for the program, the provision of more books and better school library facilities seems mandatory. Many new books and many new attractive editions of the classics and the "older" books are constantly coming on the market. However, even when money is available, it appears to be hard to get these books into school libraries. Only a small committee, meeting one or twice a year, recommends new purchases to the School Book Branch, which publishes *Invitation to Read*. Though teachers and librarians may order books from other lists, this practice is seldom followed.

As many school libraries appear to be woefully short of "newer" books, some sort of special or equalization grant seems necessary to raise the quality of some school libraries. Only by making reading as attractive as the other forms of student recreation can educators hope to have it retain its favor with high school students.

It appears that, though many Alberta teachers are adequately prepared to administer the leisure reading program, many others are deficient in book knowledge and a background of English literature.

School librarians, working in central school libraries, seem to be required in many more schools in Alberta.

As no single method of dealing with leisure reading can possibly be of much value, it can be concluded that teachers must constantly

devise methods that suit both themselves and the students they teach. *The English Journal* and other professional publications contain many hints of methodology, which are likely to be helpful to both the experienced and the inexperienced teacher.

As students frequently rely on *Invitation to Read* as a guide to book selection, one concludes that much effort should be exerted to make this pamphlet as useful as possible. It appears to require even more careful yearly revision than it now receives, and it must have better annotations.

As the students' accomplishment in leisure reading seems to have been commendable, the teachers who have guided and encouraged the students can feel that their efforts have been worthwhile.

The compilation of a guide or handbook to leisure reading as an aid to teachers and students seems desirable.

### Recommendations

The following recommendations seem to follow from the above conclusions.

1. It is recommended that five, and ultimately seven, titles per student be the minimum quota of leisure reading books in each school library. These books should be "modern", and "classics" and "school classics" in current editions.

2. It is recommended that all leisure reading teaching personnel obtain credit in two or more University English courses and/or library science courses.

3. It is recommended that many more senior high schools plan to acquire a central library under the charge of a full-time librarian.

4. It is recommended that a central library committee be instituted to work continuously on book lists, annotations, grading of books, and suggestions that will help school libraries to make leisure reading constantly more appealing.

5. It is recommended that an annual report of library facilities in each senior high school of the province be required by the Department of Education.

6. It is recommended that all schools institute or consider the inauguration of a developmental reading program for all high school grades.

7. It is recommended that no more (and no fewer) restrictions be placed upon students' leisure reading.



8. It is recommended that the leisure reading program remain "open at the top" so that able students will be encouraged to read widely, and to write and to speak about books and authors intelligently. Though more demands can always be made at all levels of student ability, teachers will have to judge the point at which excessive demands will reduce student interest, pleasure and profit from leisure reading.

9. It is recommended that a handbook to leisure reading be prepared.

10. It is recommended that the study of leisure reading be extended to include information from other Canadian provinces.

*The Editorial Committee for the*  
ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*will be pleased to consider research articles from contributors outside the Province of Alberta, for publication in the Journal.*

*Manuscripts should be addressed to the Editor.*

H. E. Smith



# The Alberta Journal of Educational Research

Vol. X, No. 2

June, 1964



THE COMMITTEE ON EDUCATIONAL RESEARCH

*Faculty of Education*

*University of Alberta*





# AN OBJECTIVE COMPARISON OF ACHIEVEMENT IN THE BASIC SUBJECTS FOR MATCHED GROUPS OF CHILDREN IN MANCHESTER, ENGLAND AND EDMONTON, ALBERTA

J. A. YOUNG

AND

M. D. JENKINSON

*University of Alberta*

The continuing criticism, which belittles education on this continent in comparison with the European, demands that objective evidence concerning achievement in the basic subjects should be obtained. Comparative studies of achievement not only serve as a firm basis for judgment, but also help to illuminate the effects of differing school organization, methods and materials upon achievement.

## **The Purpose**

The purpose of this study was to compare achievement in the basic subjects (reading, arithmetic and language) for matched groups of children in Manchester, England and Edmonton, Alberta. Comparisons were also made on the basis of sex, intellectual superiority and the specific strengths and weaknesses in achievement of two age groups in the respective cities.

## **Related Research**

No previous comparative study of the educational achievements of selected groups of children from the two nations compared in this study was located in the literature. However, a number of similar studies have compared European educational achievement particularly British with those in other geographical regions of this continent.

Scholl(4) compared the reading achievement of children of various ages in Burton-upon-Trent, Staffordshire, England, with the age norms of the Stanford Achievement Tests, which have been standardized on American children. No significant differences in the total test scores for the two groups were identified.

Preston(3) compared the reading achievement of fourth and sixth grade pupils in Wiesbaden, Germany and Philadelphia, U.S.A. The results of his study cast doubt on the validity of unqualified statements that German children read better than American children, or that the incidence of reading retardation is greater in the United States than in Germany.

Studies to compare the achievement in arithmetic of selected groups of 11-year old children in the United States and England have been completed by Buswell (1) and Tracy (5). The tests used (numerical items and word problems) produced significantly superior results for the children in the English samples. Similar results were obtained by Kramer (2) in his comparative study of children at various age levels of the elementary school.

The twelve nation study by U.N.E.S.C.O. was published in the same month in which this study was completed.

All comparative research studies that have been completed represent only a beginning in this rapidly expanding field.

### Collection of Data

#### *The Sample*

The subjects for his study were 169 nine-year old children and 173 fourteen-year old children in each city. The schools (four in each city) were selected on the recommendations of education authorities in both cities as being representative of the total school population of the chosen age groups in each city. Comparability of socio-economic status was also considered in the selection of the samples. The groups were also matched on the basis of non-verbal intelligence.

#### *The Testing Instruments*

In order to reduce the possibility that children in either city would be more familiar with the tests, both English and American tests were used. Tests used, that have been published and standardized in England were the *Schonell Graded Word Spelling Test*, the *Watts Sentence Reading Test* and the *Vernon Sentence Reading Test*. Tests used that have been published and standardized in the United States were the *California Short-Form Test of Mental Maturity*, and the complete battery of the *California Achievement Tests* at the elementary and junior high school levels.

### Treatment of Data

Mean scores and variances of the total test scores and sub-test scores were calculated and compared for both age groups in both cities. Similar calculations were made and compared for boys and girls separately at both age levels in both cities, and for those children in both cities whose I.Q. scores on the *California Short-Form Test of Mental Maturity* exceeded 130.

The significance of the differences between the mean scores for each of the groups was tested by the conventional "t" test of significance at the accepted one per cent and five per cent levels.



SUMMARY OF THE TESTS OF SIGNIFICANCE FOR THE DIFFERENCES IN THE MEAN SCORES AND VARIANCES FOR THE TOTAL GROUP AND THE SUB-GROUPS IN THE TWO CITY SAMPLES AT BOTH AGE LEVELS FOR ALL TESTS ADMINISTERED

TABLE I

	C.A.	Lang. I.Q.	Non-Lang. I.Q.	Total I.Q.	Watts	Scho-nell	Rdg. Vocab.	Rdg. Comp.	Total Rdg.	Arith. Reas.	Arith. Fumd.	Total Arith.	Mech. of Eng.	Spell-ing	Total Lang.	Total Bat-tery
Nine-Year Olds:																
Means																
Total Group Edmonton-Manchester																
Boys E <sub>9</sub> - M <sub>9</sub>	E .01	N.S.D.	N.S.D.	N.S.D.	N.S.D.	E .01	E .01	E .01	E .01	E .01	M .01	N.S.D.	E .01	E .01	E .01	E .01
Girls E <sub>9</sub> - M <sub>9</sub>	E .05	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	E .05	E .01	E .01	E .05	N.S.D.	N.S.D.	E .05	N.S.D.	E .01	E .01
Variances	E .01	N.S.D.	N.S.D.	N.S.D.	N.S.D.	E .01	E .01	E .01	E .01	E .01	M .01	N.S.D.	E .01	E .01	E .01	E .01
Total Group E <sub>9</sub> - M <sub>9</sub>																
Boys E <sub>9</sub> - M <sub>9</sub>	E .02	M .10	N.S.D.	M .02	N.S.D.	M .02	M .02	M .02	M .02	M .02	M .02	M .02	N.S.D.	M .02	M .02	M .02
Girls E <sub>9</sub> - M <sub>9</sub>	E .02	N.S.D.	M .10	M .02	N.S.D.	M .10	M .02	N.S.D.	N.S.D.	M .02	M .02	M .02	N.S.D.	M .10	M .10	M .02
Medians:	E .02	N.S.D.	N.S.D.	N.S.D.	N.S.D.	M .02	M .02	M .10	M .10	N.S.D.	M .02	M .02	N.S.D.	M .02	M .10	M .02
Gifted (over 130)																
Fourteen-Year Olds:																
Means																
Total Group E <sub>14</sub> - M <sub>14</sub>																
Boys E <sub>14</sub> - M <sub>14</sub>	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	M .01	M .05	E .01	N.S.D.	E .05	N.S.D.
Girls E <sub>14</sub> - M <sub>14</sub>	N.S.D.	E .01	N.S.D.	N.S.D.	N.S.D.	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01
Variances	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	N.S.D.	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01	E .01
Total Group E <sub>14</sub> - M <sub>14</sub>																
Boys E <sub>14</sub> - M <sub>14</sub>	E .02	M .02	N.S.D.	N.S.D.	M .10	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02
Girls E <sub>14</sub> - M <sub>14</sub>	E .02	M .10	N.S.D.	N.S.D.	N.S.D.	M .02	M .02	M .10	M .10	M .02	M .10	M .02	M .10	M .02	M .02	M .02
Medians:	E .02	M .10	N.S.D.	N.S.D.	N.S.D.	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02	M .02

M.- Significant at the indicated level in favour of Manchester.  
E - Significant at the indicated level in favour of Edmonton.  
N.S.D. - No significant difference.

The significance of the differences between the variances of the two groups was calculated and tested by the conventional "F" test of significance at the accepted two per cent and ten per cent levels.

### The Results

A summary of the results obtained is presented in the table.

#### *Differences in Mean Scores:*

*Schonell Spelling Test:* Mean scores for the Edmonton children were significantly higher than those of the comparable groups in the Manchester sample. The usual sex differences in spelling achievement with the girls' mean scores significantly exceeding those of the boys were maintained for the groups in both cities.

*California Achievement Tests:* Mean scores on the separate sub-tests were significantly higher for the Edmonton children than the Manchester children with the exception of the Arithmetic Fundamentals sub-tests.

General significant differences in the mean achievement scores in favor of the Edmonton children for the total groups in the samples from each city were maintained when the sex sub-groups were compared, but tended to disappear when only the achievement of those children in the sample from each city with I.Q. scores above 130 were compared.

*Watts and Vernon Reading Tests:* No significant differences in the mean scores of the children in the samples from the two cities were identified on those tests.

#### *Differences in the Variances:*

The variances of the achievement scores on all tests for the total groups were almost uniformly significantly greater for the Manchester groups than for the Edmonton groups.

### Conclusions

Difficulty with the format of the tests might be a contributing factor to the generally lower performance of the Manchester children. The items on the *California Achievement Tests* are of the multiple-choice type, and children in Edmonton schools have considerably greater experience with this type of test item than do children in Manchester schools. However the differences in the mean achievement scores for the two national groups do appear to be real. These differences are highly significant at both age levels, and are even more significant at the fourteen-year old level than at the nine-year old level.

The possibility that these objective tests provide a more valid measurement of the attainments of the more prescriptive and uni-



form Edmonton school curriculum than of the indefinite and varied curriculum of the Manchester schools should be considered. It should also be emphasized that these test results apply only to achievement in the three basic subjects, and might be quite different if general tests of understanding in other content subjects of the school curriculum were administered. Tests of other aspects of achievement in the three basic subjects (e.g. written language, depth and breadth in vocabulary, and the highest levels of reading comprehension) might also give quite different comparative results than those attained in this study.

Though the mean differences of scores were in favor of Edmonton, it is important to note that the differences in the variances which were statistically significant were all in favor of the Manchester groups. This suggests that the English system appears to produce greater variability in the range of achievement, perhaps because of the greater emphasis on individual differences.

Differences in promotion policies in the educational systems of the two cities also favor the children in the Edmonton sample. Promotion is automatic in Manchester schools and all nine-year old children and fourteen-year old children regardless of their achievement level were included in the sample from that city. Since comparable grade levels were chosen for the Edmonton sample, and promotion from grade level to grade level in Edmonton is partly dependent upon achievement, lowest achievers in the Edmonton schools may have been omitted from the sample due to the fact that they had not reached grade levels four and nine, from which the sample was drawn.

The importance of quick, complex mental calculations in arithmetic in English society, resulting in increased emphasis on drill in arithmetic fundamentals in Manchester schools would appear to explain the superiority of the Manchester children on this particular sub-test.

The fact that no significant differences were found in the mean achievement scores of the children on the reading tests that were published and standardized in England might be partly explained by the limited sampling of reading skills tested. Both of these tests provide a sample of the reading ability of the children on a completion type test of reading comprehension only, in a limited time. Also these two sentence reading tests appear to be less closely related to any definite school curriculum, and may therefore provide the best unbiased score for reading ability for both of the two selected national groups.

The superiority of the children in the Edmonton sample at both age levels on the *Schonell Spelling Test* appears to be related to the

marked differences in the spelling programs for the schools in the two cities. A very definite, sequential program in spelling is prescribed for the Edmonton schools with specific exercises and activities in spelling outlined for each day of the week. The Manchester spelling program represents almost a complete opposite, taking an entirely incidental approach to the teaching of spelling. Comparisons of spelling errors made in written language by children in the two cities might be quite different than those attained on this spelling dictation test.

### Implications

In general, the results of this study cast doubt on the frequently stated assumption regarding the superiority of British education over any that is given on this continent. However certain implications are suggested.

#### *(1) Curriculum and Teaching Procedures:*

The organized prescriptive program of instruction in the Edmonton schools, authorized and controlled by the Alberta Department of Education seems to produce a higher average performance rating in the three basic subjects for the majority of students than the much freer, indefinite program of instruction in the Manchester schools. However the flexible curriculum in the Manchester schools seems to widen the range of the pupils' achievement in the basic subjects significantly more than the instructional program provided for the school population of Edmonton.

These results present fundamental general questions to all educators concerned with any aspect of curriculum construction for the schools in either of the two selected countries. Is it possible to provide a school curriculum that will maintain high average performance ratings in the three basic subjects, and at the same time provide many additional opportunities for children to achieve to their maximum capacity in all aspects of the schools' program? Can official direction in matters of curriculum from centralized educational authorities produce uniform higher achievement in the basic subjects without sacrificing the provisions for individual students to proceed through the school program at their own rate?

#### *(2) Age of School Entry:*

The one year earlier starting age for most Manchester children as compared to the Edmonton starting age does not seem to be an advantage in the acquisition of skills and knowledge in the three basic subjects tested by the chosen battery of tests in this experiment. However, it does not seem fair to dismiss the importance of a planned program of learning experiences for children of chronological age-five years, simply on the basis that such a pro-



gram does not produce significant differences in measured achievement in the basic subjects at nine and fourteen years of age. Recent studies in child development have confirmed the importance of the reading readiness programs. Provisions for many experiences in oral language and for building a background of experience for children from culturally-deprived homes are two of the most essential values from such programs.

However, fundamental educational issues remain to be settled. Would the general learning and development of children throughout the schools' program be improved by definite provision for regular school attendance for all children at five years of age? If so, what types of experiences should be provided to best promote the latter development of the skills, abilities, attitudes and appreciations that are desirable outcomes of the total educational program provided by the schools? Some obvious influences on school achievement in the three basic subjects should accrue from one extra year at school if this extra year were used in the most effective possible manner.

### (3) *Streaming:*

The provision for the "streaming" of all children into separate classes, with differentiated programs of instruction according to their achievement and intellectual ability, that is common practice in Manchester schools, appears to widen the range of individual differences in achievement in the three basic subjects, measured in this experiment. The fundamental educational issue that seems apparent from the results of this study is the need for further investigation into the best administrative procedure for providing for real differentiation of the instructional program according to the abilities of the children in the schools without sacrificing a high level of general or uniform achievement. Further investigations into differentiation of the teaching methods and materials used with different groups of children as well as differentiation of the basic curriculum content are needed in both of the selected countries.

### (4) *Intellectually-Gifted Children*

The results of this experiment, indicating insufficient challenge to the intellectually-gifted students in the Edmonton sample at both age levels should cause considerable concern to educators on the North American Continent. General questions in this area with which educators in both of the selected countries are faced include the following: What additional provisions should be made in the program to meet the educational needs of intellectually-gifted children? Is it possible to meet these needs within the regular school program provided for all students, or is it necessary to segregate this particular group of children into a completely



separate school organization, in order to provide them with the best type of school program to ensure their maximum educational development?

(5) *Books and Other Teaching Materials:*

Since the results of this study show an almost uniform wider range in the achievement at both age levels of the Manchester children than the Edmonton children, the influence of the use of compulsory authorized textbooks in the three basic subjects on maintaining a greater degree of uniformity for the achievement scores of children in such a prescriptive program should be investigated. Is it possible to maintain a high mean performance in the basic subjects without restricting the core of the curriculum in the basic subjects to definite authorized textbooks? What influence does more official direction in the choice of textbooks have on increasing general achievement in the three basic subjects, while at the same time safeguarding the opportunities for a freer choice of textbooks and teaching aids for children at both of the extreme ends of the children's intellectual ability scale? If the use of prescribed textbooks contributes to the subjection of all Edmonton students to a "lock-step" curriculum, a re-examination of this prescription is very necessary.

The measurement of comparative educational achievement through objective testing has been receiving considerable attention recently. The results of this testing not only provide evidence of comparable levels of achievement but also pose many interesting questions concerning curriculum, methods and materials.

#### REFERENCES

1. Buswell, G. T. "A Comparison of Achievement in Arithmetic in England and Central California," *The Arithmetic Teacher*, 5:1-9, February, 1958.
2. Kramer, H. "Arithmetic Achievements in Iowa and the Netherlands," *Elementary School Journal*, 59:258-263, February, 1959.
3. Preston, Ralph C. "A Comparative Study of the Reading Achievement of German and American Children," *International Reading Association Conference Proceedings*, 6:109-112, 1961.
4. Scholl, Geraldine T. "The Reading and Spelling Achievement of a Group of English Children," *International Reading Association Conference Proceedings*, 6:107-109, 1961.
5. Tracy, Neal H. "A Comparison of Test Results, (North Carolina, California and England)," *The Arithmetic Teacher*, 6:199-202, October, 1959.

# A COMPARATIVE STUDY OF C. P. SNOW'S "TWO CULTURE" CONTROVERSY

DONALD F. SWIFT  
AND  
THOMAS E. LINTON

## *Introduction*

In recent years, a controversy has developed over what C. P. Snow has termed the "two culture" division in human understanding. He saw this division as a product of over-specialization within the schools and universities. This specialization started very early in England, continued throughout the individual's university training and produced two quite distinct cultures among the educated élite (1).

Professor Snow has argued that this separateness between the "arts" and the "science" culture produces a very dangerous condition in present day society. He contends that the two cultures represent different cognitive and conceptual approaches to human problems. The individual with an arts background is generally unable to understand even the most rudimentary scientific developments. His training has promoted a very limited awareness of modern technical and scientific discoveries.

This lack of understanding tends to be encouraged by the attitudes which an "arts" man learns during his university study. The same may be said of the "science" major in that he frequently knows little about the "arts", and has prejudiced feelings towards people majoring in the arts areas and their disciplines.

These attitudes are promoted by the divisions of interest which are present in various university departments. Not only is there a dichotomy between the "arts" and the "sciences", but there are important differences between disciplines within a particular field. These differences may be emphasized by the students' professor and by the more advanced students in a discipline who informally make sure that the prejudices and partisan values of their field are accepted by the novitiate (2).

The central problem is one of communication between those who have specialized in the Arts and those majoring in the sciences. Both of these individuals may move towards leadership positions in their societies, and be seriously handicapped in their ability to understand the ideas and values of the other individual. At the present time many of the most important decisions affecting the survival or destruction of human life are made by the scientist. The arts man frequently is the administrator who smooths the path for



the scientific act. He seldom understands the technical ramifications of the orders he expedites. His questions about the meaning of a particular move are usually viewed by the scientifically trained staff member as peripheral to the problem itself, if not as essentially unrelated to the issue as perceived by the technician. This further lessens the possibility of good communication between the scientist and the "arts" administrator (3). It also serves to increase the social and political power of the technical specialist, and to lessen the influence of the nonspecialist.

Historically the liberal arts were in the favored position. The cultured and successful individual had only to major in the arts and evidence little interest in the sciences. European tradition still tends to place greatest emphasis on the cultured person as one who knows literature, history and philosophy, and has an active interest in the fine arts. Technical and scientific training were not traditionally endowed with the same prestigious aura as were the so-called "Liberalizing" arts (4).

In academic circles it is traditionally assumed that an "arts" background is essential to liberate an individual from his personal and parochial past. However, studies of this assumption have indicated that college training may serve to reinforce old prejudices, and present the student with a more sophisticated means of maintaining his old interests (5).

### *Method*

In an attempt to examine C. P. Snow's hypothesis about the "Two Culture" controversy, Kenneth Richmond in England has developed a Culture Test (6). He administered the test to grammar and public school boys in the sixth form, to students in the university education department, and to students in women's training colleges. He thus provided a series of average scores indicating certain levels of attainment for different student bodies in England. The present authors thought it would be interesting to give the same test in similar educational institutions in Canada and the United States and compare these to the scores achieved in the English schools.

It is difficult to compare educational institutions in different cultures. For the effects of the culture itself as well as the many contrasts in academic approaches make these comparisons artificial. Granting these conditions the authors selected the following institutions:

1. A Provincial University, recently established in one of the larger cities (270,000) in the prairie provinces of Canada.
2. A California "State College of Applied Arts and Science" which has one of the larger teacher training divisions in the states. The college is located in a city of over 3,000,00 people.

3. A major State University located in one of America's larger Eastern seaboard cities (population over 6,000,000).

In general the comparisons were focussed on teachers in training, although additional comparisons were made between teaching majors and non-teaching majors in the Canadian University.

The culture test was composed of two parts with twenty questions in each part. One part was made up of general questions from the arts, while the second part contained more specific questions from the sciences. Mr. Richmond assumed that the mean scores of different groups on the arts and science parts of the test would be an indication of their general level of knowledge in these respective fields.

This article is not presented as a critique of the assumptions of C. P. Snow for it is felt that a different kind of instrument needs to be developed for that purpose. However, it was felt that the Culture Test was indicative of certain features of British secondary education, and in that sense a general comparison with Canadian and American students would be worth presenting.\*

The Culture Test was administered to 454 first year students, 47 second year students, and 79 fourth year students in the Canadian Provincial University. In addition 165 students in Education were tested at the California State College, and 111 Education majors attending the American State University.

TABLE I  
ENGLAND

	Arts	Science	Total
Public School Sixth Form (17-18 years) .....	7.3	10	17.3
Grammar School Sixth Form (17-18 years) .....	5.4	10.9	16.3
Two University Education Departments { .....	7.4	8.6	16
	6.8	9.3	16.1
Three Women's Training Colleges { .....	4.4	5.6	10
	4.7	6.0	10.7
	3.2	3.7	6.9

\*Because the Culture Test used in this research was not scientifically constructed and validated, the authors felt that statistical testing for the significance of the differences with the mean scores was not justified. All remarks regarding differences, therefore, are impressionistic.

### *Results*

The results are presented as the average number of questions correctly answered by a particular group of students. The arts and science average scores for each group are reported separately and then as a total out of a possible maximum of forty questions in the test. The tables that follow indicate the scores on the Culture Test obtained in (1) England, (2) Canada, and (3) America. After each of the tables there is brief discussion of the possible implications of the scores.\*

### *Discussion*

The Women's Training Colleges all received low scores. These were not university students, the large majority of them are not qualified for university entrance. However, their training does suit them for teaching in the elementary schools. This would be a further indication that academically weaker students frequently enter the teaching profession as elementary school teachers.

We see that the Sixth Form scores are similar to those obtained in a University Department of Education which could mean that the British university students (especially in science) tend not to go into teaching. There seems to be less evidence for this in Canada and the United States. It would be expected that the Public School students would receive higher scores. They tend to come from more successful business and professional families, and are a highly selected group. Their schooling is primarily university committed, and the tone of school life emphasizes the importance of humanistic and scientific learning. The grammar school group represents a lower socio-economic level than the public school group, but there may be more high achieving individuals among these students. For many "average home" students gain access to the grammar schools on the basis of their eleven plus examinations and academic record. This would mean that the grammar school is viewed as a primary ladder of social mobility for the high achieving less affluent student (7).

One would expect the University Education Departments to do better than the Public and Grammar School students. The University students are older and have finished their undergraduate training. They have experienced the many cultural advantages associated with student life. The differences between the University Education Departments and the Public and Grammar School scores may be viewed in these terms. The Public Schools have traditionally been training centers for the children of the élitist

---

\*The material in this study is presented as an interesting note on education in three societies. It does not pretend that these figures may be used as valid reflections on the educational efforts in these countries. A more rigorous sampling design and testing approach would be required if a serious attempt at cross cultural evaluation were to be made.



families in British society. They represent social as well as academic standards not within realistic reach of the average family. It is obvious that the son of an "establishment" family has many cultural advantages not present in the background of children who enter the Grammar Schools, or those who choose to become teachers.

This may account for the difference between the Grammar School arts score and the Public School score. Further the teachers in training in the University achieve about the same level as the Public School students, even though they are older and have had more academic experience. The Public Schools, the Grammar Schools and the University Education students have similar science scores, probably because this involves more straight forward retention of "school taught" material. However, it should be noted that the University students are again behind the Public and Grammar school students. Is this an indication that British teachers in training are derived from a lower achieving academic group?

TABLE II  
CANADA

Western Canadian Provincial University	Arts	Science	Total	N
1st year Education .....	2.11	7.40	9.51	65
1st year Engineering .....	1.95	8.91	10.90	91
1st year Chemistry .....	2.07	9.42	11.50	114
1st year English .....	2.20	7.09	9.29	100
1st year Commerce .....	1.62	8.62	10.22	84
1st year Average .....	1.97	7.77	9.74	454
2nd year Education (Jr. E.) .....	2.64	7.8	10.44	47
B.Ed. Fourth Year .....	3.87	11.33	15.2	79

\*The five different departmental areas making up the first year average are not mutually exclusive categories. Many of the students taking the Chemistry, English and Commerce first year courses may also become teachers. The groups making up the "First Year Average" represents a near total sample of the first year university class (N = 454).

*Discussion*

The scores of the Canadian first year students do not differ markedly from one department to another. Most significant are the very low scores on the arts questions for all the first year students. One example of this may be seen when a single group is examined. Among the education students there were twelve who scored zero out of the twenty questions, and another thirteen who obtained one out of the twenty for the arts questions. The second year education students showed little improvement, while the fourth year teaching majors evidenced some improvement, but were still very low in the arts area.

What do these low scores on the arts questions represent? The inability to answer more than one or two of the twenty items may indicate a cultural and educational deficiency in the students' background. Even when one grants that the arts questions may be somewhat esoteric, the very low scores are difficult to interpret.

Most of the arts questions are based on information which would be found in the literary and cultural pages of the better magazines and newspapers. Further they would be based on a general level of cultural and literary knowledge which is more characteristic of certain urban centers than of others. In this sense one would expect that the university students in Montreal and Toronto would have higher mean scores on the arts questions than similar students from the universities in the prairie provinces.

It may be argued that a low arts score is not a serious matter for the non-education major. But when one realizes that many of the education majors will enter the teaching profession at the end of their second university year, the situation is worthy of notice. It is only a very recent development that two years are required before commencing to teach in the public schools in this province. Previously one year of preparation for teaching was sufficient, and a large number of teachers in the schools began with this level of training.

When one considers the small amount of education the teachers have experienced, and the "cultural" level of development in the region, the low arts scores are more readily understood. The question that results is what will these teachers be able to pass along to their students in the way of cultural information?

The difference between the scores of the first and fourth year education majors seems to provide some argument in favor of a four year preparation of teachers. However, the small amount of improvement on the arts questions would not indicate much of a return for three more years of university study. The small improvement may indicate something about the "cultural" life in the university apart from the formal instruction in the classroom.

On the positive side the first year students have been well prepared in the “science” area. They have evidently recalled the factual material presented in their high school science classes.

In considering the fourth year B.Ed. performance the same trends are noted. The arts scores are very low for individuals who have been involved in university life for four years.\* It may be questioned whether the intensive cramming of material for a large number of courses in seven month sessions is the most beneficial approach which can be offered the intending teacher?

But on the science questions the fourth year Canadian students did better than any of the other schools tested. This is partly a result of the fact that many of these students were science majors. Here the substantial gap between the first year and fourth year promises a valuable increment in scientific knowledge for the student.

TABLE III  
AMERICA

	Arts	Science	Total	N
1. California State College				
Education—Undergraduate .....	5.10	7.50	12.60	72
Education—Graduate Students .....	6.10	8.81	14.91	93
2. Eastern State University				
Education—Undergraduate .....	5.25	8.12	13.37	51
Education—Graduate Students .....	7.07	10.98	18.05	60

Discussion

The two colleges represented in Table III are similar in certain ways. They both cater to an urban student, many of whom work on a full or part time basis in addition to their university studies. A large number of the students at both colleges come from minority groups who attend the low cost state subsidized schools as a means of improving their social and economic position. Further there is a large amount of “cultural” activity in these cities, and it would be difficult for a socially mobile person to avoid even casual contact

\*These are generally not students who work apart from their university studies. Their leisure is not committed to economic responsibilities as is the case for many of the students in the urban colleges tested in the U.S.



with cultural information. There is literally an endless variety of entertainment at all cultural levels available without effort to the public. There is an abundance of consumer goods, recreational alternatives, and varieties of living styles within reach of many of these students.

These living conditions are different from those present in the western Canadian University where social and cultural limitations are more obvious. Considering the level of the culture in these American cities the scores on the Arts questions are not very impressive. The graduate student total score at the Eastern University is a good deal higher than the total scores of the other groups. This may represent selection factors in terms of who goes on to graduate education. The graduate students are a higher achieving group in the arts and science areas as well. While this difference may be a function of an inadequate sample, it may also indicate that the graduate students are more knowledgeable at the one institution.

Also the eastern university students are slightly higher on all comparable scores than the California State College. The eastern school is usually viewed as a more competitive college in terms of types and numbers of students seeking graduation requirements. However the California State College is a newer school and has served as a senior college for most of its existence. This means that the junior colleges in the area have channelled their students into this school.

So far the scores have been dealt with according to specific countries. The most important conclusions will be derived from a cross-cultural comparison of scores in different societies. If all of the arts scores are placed together on one table and the science scores on another an overview of the relative achievements of the different schools is gained. Tables IV and V serve that purpose.

### *Discussion*

At the top of the Arts Table we find the English University and the Public School. While at the lower end of the table are the Canadian first year University students. It seems reasonable to suggest that the mean scores reflect the level of cultural life within a particular college.

The mean scores may reflect the degree of urbanization of the institution in which the "Culture Test" was administered. The more rural locations offer less in the way of general cultural advantages. Further the whole approach to the "socialization" of the child differs in its lessened emphasis on the value of "arts" knowledge. Provincial values tend to stress practical information and a more vocational approach to the meaning of education. However,

we are less concerned with the explanation for the cultural level than with establishing the presence of the differences.

When the "science" scores are placed on a similar table we see that there are some changes in the position of the schools.

TABLE IV  
MEAN "ARTS" SCORES OF GROUPS WITHIN THE  
DIFFERENT INSTITUTIONS

English University Department of Education .....	7.40
English Public School Sixth Form .....	7.30
American University Graduate Education .....	7.07
English University Department of Education .....	6.80
California State College—Graduate Education .....	6.10
English Grammar School Sixth Form (16-18 years old) .....	5.40
American University Undergraduate—Education .....	5.25
California State College—Juniors (Third Year) .....	5.10
English Women's Training College .....	4.70
English Women's Training College .....	4.40
Western Canadian B.Ed. Final and Fourth Year .....	3.87
English Women's Training College .....	3.20
Western Canadian 2nd Year Education .....	2.64
Western Canadian "1st Year Average" .....	1.97

TABLE V  
MEAN "SCIENCE" SCORES OF GROUPS WITHIN THE  
DIFFERENT INSTITUTIONS

Western Canadian University B.Ed. Final Year .....	11.33
American University—Graduate Education .....	10.98
English Grammar School Sixth Form .....	10.90
English Public School Sixth Form .....	10.00
English University Education Department .....	9.30
California State College—Graduate Education .....	8.81
English University Education Department .....	8.60
American University—Undergraduate Education .....	8.12
Western Canadian 2nd Year Education .....	7.80
Western Canadian 1st Year Average .....	7.77
California State College—Junior Education .....	7.50
English Women's Training College .....	6.00
English Women's Training College .....	5.60
English Women's Training College .....	3.70

*Discussion*

There is no real difference between comparable Canadian and American students, while intending teachers from both countries have better scores than the women in the three English Training Colleges.

The very low scores at one of the English Training Colleges would indicate that these students are achieving at much lower levels than teacher trainees elsewhere. These students will probably be very limited in their ability to enhance the "science" knowledge of their pupils. One wonders about the criterion for selection in these institutions, and about the kinds of schools in which these teachers will work. Are these the elementary teachers who find positions in the poorer socio-economic school districts (8)? Will their employment tend to limit the social and intellectual growth of their students?

An interesting difference between the "science" and "arts" scores is the much higher score gained by the Canadian Bachelor of Education student. He moves up from a very low "arts" score to the top position on the table of "science" scores. This represents three years of University study and some improvement over his first year "science" score. It may further indicate the value placed on scientific and useful vocational information which seems to be present in this institution. Students appear to view education as a functional process, useful for its value in preparing the individual for his vocational future.

*Summary and Conclusion*

A "Culture Test" developed by Kenneth Richmond in England was given to students in a Western Canadian University, a California State College and an Eastern American University. The aim of the "Culture Test" was to evaluate C. P. Snow's hypothesis about the conflict between the "Arts" and the "Science" culture.

This article was an attempt to compare educational institutions involved in the teacher training process, as well as comparing intending teachers to English Public and Grammar School students. The study did not attempt to measure C. P. Snow's ideas, but rather to utilize the "Culture Test" as a means of comparing levels of general knowledge in differing societies. Certain tentative conclusions may be drawn from these figures.

1. In recent years the importance of scientific training has been stressed in the schools. These tests results indicate that in Canada and America the students scored better on the science questions than on the arts questions. The differences may be caused by the popularity gained by science in the last decade. More scientific "facts" have entered into the teaching of science, and many new



courses and interesting approaches have been initiated. The traditional dominance of the "Arts" culture has been lessened and perhaps placed in a secondary position. It would appear that the "Arts" culture is in a very subordinate position in the Western Canadian area that was sampled in this study. This condition will probably continue unless the schools make a serious effort to improve this deficiency.

2. The schools differed along certain dimensions, namely whether they were rural or urban, and whether they were mass or élitist in their selection policy towards students. The English Public and Grammar Schools represented an élitist educational approach and their scores were among the highest. While the California College has more of a mass approach and their scores were poorer than the English schools.

A rural versus urban factor may also have been present in the very low "arts" scores obtained by the first and second year Canadian University students and the students in the Women's Training College in England.

3. There was a difference between teaching majors and non-teaching majors. Some of the lowest scores were obtained by the English women in the training colleges. In contrast the English Public School students, who do not generally become teachers, had higher scores. The Canadian "arts" scores were very low for individuals who will shortly serve as teachers in the public schools. The implication is simply that these individuals who are to serve as transmitters of the cultural heritage should be provided with broader awareness of the "arts culture".

Finally the authors wish to state that the sampling procedures, as well as the "Culture Test" utilized in this study, do not represent a carefully designed and controlled study. Further research is being conducted by Professor Richmond; and by the present authors toward that goal. This study should be viewed as a simple presentation of the mean scores obtained on Professor Richmond's "Culture Test" in somewhat comparable institutions. What these scores actually indicate is difficult to say unless a large number of variables are considered in the different societies (9).

#### REFERENCES

1. Snow, C. P., *Two Cultures and the Scientific Revolution*, Cambridge University Press, N.Y., 1959.
2. Riesman, David, *Constraint and Variety in American Education*, Doubleday Anchor Books, Garden City, N.Y., 1958, p. 79.
3. Richmond, Kenneth, "Two Culture Tests", Off Print from *New Saltire* No. 6, December 1962, p. 3.
4. Butts, R. Freeman, *A Cultural History of Education*, McGraw-Hill, N.Y. 1947, p. 571.

5. Key. V. O., Jr., *Public Opinion and American Democracy*, Alfred A. Knopf, N.Y., 1961, p. 343.

"In the largest sense the educational system is a great mechanism for conserving the values, the institutions, the practices of the political order, as it inculcates from the primary school even through college the memories, the unities and the norms of the system."

6. Richmond, Kenneth, "A Measure of Culture", *The Times Educational Supplement*, Friday, September 29, 1961, p. 381.

Examples of questions from the Arts and Science areas used by Mr. Richmond in his culture tests:

1. *Arts.*

The characters in "Under Milkwood" were—

(a) English, (b) Breton, (c) Scots, (d) Welsh, (e) Irish, (f) French-Canadian, (g) none of these.

2. *Science.*

Which of these would you expect to be most interested in the mitochondrion—

(a) a classisist, (b) an astronomer, (c) a philologist, (d) a physiologist, (e) an engineer, (f) none of these.

7. Halsey, A. H., Floud, Jean, Anderson, C. Arnold, *Education, Economy and Society*, The Free Press of Glencoe, 1961, p. 154-157.

8. *Ibid*, p. 539-543.

9. Richard, Kenneth, "Two Culture Tests", Off Print from *New Saltire* No. 6, December 1962.

In this regard Mr. Richmond has said that: "In schools and colleges the hoohead over the two cultures has given rise to various makeshift proposals in which 'breadth', balance, and all 'roundness' figure as catchwords. All kinds of airy presuppositions are being made. Not least among them is the assumption that a broad based curriculum . . . is an essential prerequisite for breadth of mind in the learner . . . It was against this confused background that the idea of pioneering "Culture Tests" began to take root in my mind. . . . It seemed that an attempt had to be made to devise a yardstick of sorts, and to use it to obtain solid data in a field which was otherwise all too hazy. Inevitably, the tests so far devised have been crude affairs."

# A FACTORIAL STUDY OF STUDENT ASSESSMENTS OF TEACHER PERFORMANCE\*

B. E. J. McBRIDE  
*Institute of Education*  
*MacDonald College*

## The Problem

The evaluation of teaching proficiency has been the subject of intensive but unsystematic attack since Book's original study in 1905. As Anderson and Hunka (1963) have reported, the inter-rater reliability of teaching performance is low and, therefore, the value of say, Faculty of Education practice teaching ratings is dubious.

## Theoretical Considerations

### *The Good Teacher: A Category*

The problem of evaluating teaching performance is essentially one of identifying the good teacher. Categorization, as explained by Bruner (1951), is directly applicable to this problem, the concept, "good teacher" being a category, the identifying or criterial attributes of which must be based on value judgments. Since systems of values vary from individual to individual, the category "good teacher" must necessarily be a disjunctive one. (Bruner, 1956) In a disjunctive category there is no core of universal common factors; in fact, the identifying attributes are likely to bear little similarity to one another and category members to differ markedly. The consequences of the persistent but unsuccessful effort to change the category "good teacher" to a conjunctive one, wherein the criterial attributes are well known and undisputed, are serious. Researchers proceed as if there were a concept 'a good teacher' whose criterial attributes are known to all and manifested in every teaching situation. This is not the case.

We might inquire what determines that which is characterized as good teaching? How does one arrive at a definition of good teaching? As Bruner relates, perception occurs on the basis of some expectancy or an hypothesis, nebulous though it may be. We not only see but we look for, not only hear but listen to. In short, perceiving takes place in a "tuned organism". (Bruner, 1951, pp. 123-124) The tuning is largely a product of past experience and training. Conceptual systems, described by Harvey, Hunt and Schroder (1961, p. 204) as the "immediate antecedents of cognitive

---

\*The author is indebted to Dr. C. C. Anderson, under whose supervision the study was conducted.



and behavioral outcomes," develop, stimuli being perceived and interpreted as either refuting or confirming these conceptual systems.

The evaluator's conceptual system determines which aspects of a teaching performance he perceives as being the criterial attributes of "good teaching". He has, in effect, a set of established standards to which he believes teaching must conform. In an effort to preserve this conceptual system, he attends to aspects of input which reinforce it and screens out dissonant information. This is particularly true of people with closed minds as Rokeach (1960) suggests.

The basic characteristic that defines the extent to which a person's system is open or closed, is namely, the extent to which the person can receive, evaluate, and act on relevant information received from the outside on its own intrinsic merits, unencumbered by irrelevant factors in the situation arising within the person or from the outside. (1960, p. 57)

It seems reasonable to assume that an adult will be more susceptible to such close-mindedness than the child, who, in his lesser sensitivity to the socially desirable attributes of perception, should have a more flexible value system. The child's evaluation of teaching, therefore, will be less likely to be affected by defensive distortion, and accordingly it merits research attention.

### *Student Evaluations of Teaching Performance*

The opinion that student evaluation of teacher performance deserve attention is by no means new. One need only consider that the student is a 'constant consumer'; that his judgment is not based on the somewhat false situation created when a supervisor enters the classroom, and that his evaluation is tempered by those of other students experiencing the same teaching, to recognize the likelihood of such ratings' being credible sources of information. The reliability coefficients attributed student ratings vary from .55 to .95. (Guthrie, 1927; Boardman, 1930; Root, 1931; Wilson, 1932; Bowman, 1934)

Student evaluation of teaching is not, however, without opposition. The student, it is pointed out, is not trained to observe teaching (perhaps this is the lack of bias valued by supporters of student rating). It is suspected that he may be influenced by the degree of his own success, that his sex may determine the favorability of his attitude towards teachers of one sex or the other, and that his age may affect his ability to rate. Also, it is asserted that the students' rating teaching is damaging to teacher morale and to student-teacher relations. Most of these charges have been investigated. Student success has been demonstrated to bear no relationship to student evaluation of teaching proficiency, (Remmer, 1930; Bowman, 1934) and, student rating, rather than damaging teacher morale, has been found to promote better teaching and

increased teacher self-understanding. (Bowman, 1934; Ward, Remmer and Schmalzried, 1941; Bryan, 1942; Amatora, 1951; Savage, 1959) Factors such as the age and sex of student raters, however, are considered more likely to affect ratings, although to a limited degree (Davenport, 1944). Another study (Morsh and Wilder, 1954), reports that the student factors, class size, sex, age, intelligence and mental age have little bearing on the evaluation made by students.

### **The Relationships Between Student Evaluation and Other Criteria of Teacher Proficiency**

The findings regarding the efficiency with which students rate teaching are indeed diverse. An attempt to validate these ratings might be made through comparing them with the following criteria of teaching effectiveness, each of which is considered in turn: pupil growth scores, supervisory ratings, the accuracy of teacher perception of student body opinion, student rating of classroom interaction, and teacher self-ratings.

Pupil growth is perhaps the most widely recognized criterion of good teaching. Even so, limitations are many, for such growth may not be entirely the product of teacher effect though pupil intelligence and initial achievement be partialled out. Other factors remain uncontrolled. (Barr, 1953, p. 8) Some lie within the pupil in the form of general and specific abilities, previous training and motivation; others such as socio-economic status, respect for education and assistance by family members are aspects of the home environment; and still others like morale, physical facilities and the influence of other teachers might be attributed to the school and community. The difficulty of measuring pupil growth in its fullest sense, encompassing the development of personality, mental health, emotional maturity, and aesthetic interests, further complicates the problem. The few studies based on the pupil growth criterion present conflicting findings, some (Boardman, 1930; Lins, 1946; McCall, 1952; Guthrie, 1953; Keller and Clark, 1954) apparently laying a foundation for a relationship between student ratings of teaching and pupil growth scores, others (Jensen, 1961; Schmid, 1961) asserting that pupil gain bears no relationship to rating factors. Further study of pupil rating and the use of improved growth measures might lead to the solution of the problem.

The correlations between supervisory ratings and student ratings of teacher performance are frequently significant and positive as illustrated in the accompanying table. (Table I) A more recent investigation carried out in this province (Anderson, 1962) reveals contingency coefficients significant at the .05 and 0.1 levels respectively between pupil and supervisory and pupil and Faculty of Education raters. (Table II)

TABLE I  
CORRELATION OF STUDENT RATING WITH ADMINISTRATIVE RATING OF  
INSTRUCTOR EFFECTIVENESS

Author	No. of Teachers	Year	Level	Type of Relationship	Degree of Relationship
Knight	39	1922	elem. and high school	supervisory rating and student ranking	.74
Boardman	88	1928	high school	supervisory and student rankings	.56
Light	28	1930	high school	supervisory rating and student ranking	"close agreement"
Greene	32	1933	college	ratings by dean, peers, and students	
Bowman	30	1934	student	critic teachers' and student ratings	All but 1 of dean's first 10 and 2 of peers' first 10 same as students, first 10 —.50 to .47
Starrak	entire faculty	1934	college	educational specialist and student graphic rating	
Bryan	22	1937	senior high	administrative and student ratings	.36
	41		junior high	administrative and student ratings	.53
	22		senior high	administrative and student ratings	.36
	41		junior high	administrative and student ratings	.55
Brookover	33	1940	high school	supervisory and pupil ratings	— .08
	12		high school	av. 2 supervisory ratings and pupils ratings	.63
Ward, Remmers	40	1941	student	supervisory and student ratings	.875
Schmalzried	27	1942	student	supervisory and student ratings	
Porter	66	1945	high school	supervisory and student ratings	
Brookover	all				"close agreement"
Von	male				"slight positive relationship"
	50	1946	high school	educational specialists' rating based on interview and autobiography and student ranking	none of 27 r's significant at .01 level. Range —.06 —.30, 28.
Lins	58	1946	high school	composite supervisory rating and student ranking	
	all				
	female				



TABLE II  
CONTINGENCY COEFFICENTS AMONG VARIOUS  
CRITERIAL INDICES OF TEACHING PROFICIENCY

Pupil	AB	Sup.	Pr. T.	Univ. Ave.
Pupil .....	.59**	.33*	.40**	.24
AB .....		.33*	.18	.17
Sup. ....			.21	.15
Pr. T. ....				.15

Pr. T.=Practice teaching marks given to the students while they were in training in the Faculty of Education in Edmonton, (1958-1959).

Univ. Ave.=Average mark in all other courses in the Faculty of Education in 1958-1959.

AB=In the field ratings by Anderson-Black-Buxton in the spring, 1960, after the students had almost completed one year of teaching.

Pupil=Questionnaire ratings by students of these teachers (grades 4-9) gathered by Anderson-Black-Buxton during their visit to the teacher and obtained while the teacher was outside the room.

Sup.=Superintendents' estimates gathered in the spring-summer of 1960.

\*Significant at the .05 level of confidence.

\*\*Significant at the .01 level of confidence.

1 Reproduced from Anderson, C. C., (1962)

Similarly, the findings with regard to the relationships between teacher perception of student opinion or affect and student ratings of teaching tend to support student ratings. Indeed, the most frequently mentioned characteristic of the ideal teacher is possibly the possession of a keen understanding of students. The study which lends most direct support to this point of view is that by Gage and Suci (1961), in which a significant relationship was demonstrated between the degree to which high school teachers were able to estimate percentages of various student responses to a number of controversial questions and the 'positive affect' elicited by these same teachers in the student body. Further support is found in the work of Gage, Leavitt and Stone (1955), and to a certain degree in studies of leadership, the more successful leader (here, the teacher) being better able to predict the feelings and consequent action of group members and accordingly adjust his or her own behavior so as to produce harmony and direct positive affect towards himself. Steiner (1955) suggests that the leader, on the contrary, need be concerned with the hierarchy of the organization rather than individual idiosyncracies. Also, one must recognize that possible accurate leader perception of the feelings and attitudes of others and the favorable impressions of the perceptive leader on the part of his followers may both be the products of a third influence.

Research attention has recently been directed to the interaction between teachers and pupils. Drawthorne (1954), one of the few to explore the relationship between pupil ratings of teacher-pupils interaction and pupil ratings of teacher effectiveness, reports a correlation of .557 between the two; while Brookover (1940), as the result of an earlier and similar study reported a corresponding correlation of .639. A recent study by Gold (1962) related teacher effectiveness to the number of isolates in the class, teachers revealing a greater degree of dominative type thinking (measured on the F scale), being found to have a significantly greater number of isolates in their classes. Gold's work bears some similarity to that of a number of others, all of which lend it support. (Anderson, 1939, 1943; Anderson and Brewer, 1945, 1946; Anderson, Brewer and Reed, 1946; Thelen, 1957, 1959; Cogan, 1956, and Flanders, 1958, 1959, 1961a, and 1961b)

Self-ratings, subject to all the limitations of introspection, are not often recognized as being of much importance in the evaluation of teacher performance. A study by Webb and Nolan (1955), in which a correlation of .55 is reported between self-ratings and pupil ratings, suggests that they might be considered to lend some support to the use of student observations as a means of evaluating teaching proficiency.

It would appear, thus, that pupil evaluations of teaching performance possess sufficient merit to warrant further investigation, and on the basis of this tentative conclusion, the study described in subsequent sections was carried out.

### **The Development of Hypotheses**

As Gombrich (1960) so aptly states, evaluative reactions . . . testify to the constant scrutiny with which we scan our environment with one vital question—are you friendly or hostile, a 'good thing' or a 'bad thing'? (1960, p. 232) In evaluating input, the individual must focus on certain cues, (Berlyne's (1960) stimulus selection). These cues are therefore of immediate concern in any investigation of teaching. To what cues does the student attend?

Studies indicate that students are keenly aware of sincere, understanding and responsive behavior on the part of teachers. (Cotsonas and Kaiser, 1962; Gage and Suci, 1951; Gage, Leavitt and Stone, 1955; Flanders, 1958, 1959, 1961b, and Gold, 1962) The student can then, in his responses to questions pertaining to classroom proceedings and his own related attitudes, be expected to reveal an awareness of, and concern with teacher-pupil rapport, particularly in terms of freedom to actively participate in classroom activity and discussion. Attitudes toward such freedom might be expected to be generally favorable, since pupil participation is one



of the marks of the classroom in which the teacher reveals concern with the individual student and understanding of his needs.

Just as he is aware of his freedom to participate, so also the child may be aware of and respond to the constraining, disciplinary and restrictive aspects of his environment. Both at home and at school, disciplinary measures serve as directives toward what is considered to be desirable behavior.

The control of behavior by means of reward and punishment or approval and disapproval can be explained in terms of arousal induction and reduction. As Berlyne (1960) relates, the individual is constantly seeking a steady state of optimum arousal, (pp. 200-209), significant departure from which represents 'unpleasure'. Both internal and external cues continually disturb this balance. In an effort to produce certain behavior, significant others induce arousal and set up corresponding incentives, the reaching of which by suitable behavior may reduce arousal, thus establishing working habits. Other theorists also explain the establishment of behavior patterns through the use of punishment and reward or anticipation of either. (Mowrer, 1950; Davis, 1948; McClelland, 1961)

The following hypothesis, with respect to the dimensions of student response to classroom events, ensue: that factor analysis of student responses will reveal

1. A factor describing the extent of student participation in classroom activity, and
2. a factor describing student awareness of and concern with the disciplinary and restrictive facets of the classroom,
3. a third hypothesis, or rather likely experimental possibility would be concerned with the possibility of developmental changes in the strength and incidence of these factors.

### The Procedure

As part of a larger teacher evaluation program carried out in the spring of 1960 by members of the Faculty of Education of the University of Alberta, student responses to twenty questions regarding classroom activity, impressions of classroom interaction and interpersonal relationships, and attitudes toward work and attendance were collected in the classrooms of teachers who had been enrolled in the Junior Elementary and Intermediate classes in the Faculty of Education at the University of Alberta in the 1958-1959 winter session. All the teachers in whose classrooms student responses were recorded had thus had the same amount of teaching experience and had experienced the same training. The responses of 2023 students from grades three to nine inclusive were factor analyzed to determine whether the variations revealed in twenty-one responses could be accounted for in terms of fewer categories.



In each grade and in groupings from multiple-grade classrooms, the factors isolated were studied with reference to the items contributing most to them, and, on this basis they were matched with similar ones from other groups and grades. Tucker's coefficient of congruence (Harman, 1960, pp. 257-259) was employed to determine the degree of relationship between each factor in each grade or group and the most similar one in each other grade or group.

### **The Findings**

Five factors were found to be constant in all or most of the groups. They are discussed in order of importance in succeeding paragraphs.

The first of these factors is a general satisfaction factor indicative of approval of the class and its activities and individual pleasure at participating therein. The most important items in terms of the size of their loadings and thus their contribution to this factor: Do you like to be in this class? Are you proud to be in this class? The measure of positive affect might be compared with the latter of two dimensions advanced by Foa (1961) as explaining human relationships, that of love-hostility and tending, of course, in the direction on the continuum labelled "love". It is not unlike what Lawrence (1961, p. 196) refers to as spontaneous informal group structure and Anderson (1939) labels socially integrative behavior.

The second factor mirrors the restrictive or disciplinary aspects of the classroom and the typical problems of noise, time wastage, and rule-keeping along with the prevalence of teacher scolding. The following items are examples of those contributing to it: Do the pupils often make so much noise that it's hard to work? Do some pupils break class rules a lot?

The remaining three factors are not so clearly defined as the former two; nor do they appear in all the grades and groups.

One of these, the third factor, might best be described as a participation factor, the students in it having revealed concern with the extent to which they work in groups, plan activities and formulate rules to be obeyed. These items serve to define it: Did the pupils make up the class rules they are supposed to obey? Do you help plan what the class is going to do?

Factor four reveals a tendency for the student to question the instruction offered coupled with a desire to absent himself from it, as revealed by the following items: Do you ever feel like staying away from school? Do you ever wonder why you have to do what you are doing in school?

The fifth factor relates to the individual's doing his best or doing satisfactory work and is well illustrated in items as follows: Is your

work usually good enough? Do you always do your best in this class?

On the basis of these findings, the first hypothesis, that factor analysis of student responses to classroom events would reveal a factor describing the extent of student participation in classroom activity, receives some support. The various factors chosen as representing participation in the groups compared fail, however, to relate to one another sufficiently to warrant one's concluding that students in various grades respond consistently. Further research with more refined and elaborate measures of participation might locate and define this quality more exactly.

The second hypothesis, that an analysis of student responses to classroom proceedings would reveal awareness of and concern with the disciplinary and restrictive facets of the classroom is also supported, the factor being relatively stable, students having responded consistently in terms of their cognizance of this aspect of the classroom.

The third hypothesis relating to the possibility of there being developmental changes in the strength and incidence of factors receives little support in the form of clear evidence, although the following two trends emerge:

- a. a relationship between sex, dissatisfaction with classroom instruction and consciously doing one's best work at school. Whereas boys tend to be dissatisfied to a greater extent and desirous of avoiding the disliked instruction, girls appear to continue to do their best work and try to please teachers, even in the later years of junior high school.
- b. a tendency for students to become, with increasing age, to a greater extent group-oriented, older students expressing pride in classes to which they belong, whereas younger ones place greater emphasis on individual relationships.

In this study ground is broken for a more intense study of student responses, the dominant dimensions of student response having been more effectively established than hitherto.

#### REFERENCES

- Amatora, Sister M. Diagnostic teacher rating. *Catholic Ed. Rev.*, 1951, 49, 675-680.
- Anderson, C. C. The evaluation of teaching proficiency: research findings and theoretical considerations. Unpublished manuscript, 1962.
- Anderson, C. C. and Hunka, S. M. Teacher evaluation: some problems and a proposal. *Harv. Ed. Rev.*, 1963, 33, 74-96.
- Anderson, H. H. Domination and socially integrative behavior. In Roger G. Baker, Jacob S. Kounin, and Herbert F. Wright (eds.) *Child behavior and development*, New York: McGraw-Mill, 1954, 459-483.
- Anderson, H. H. Measurement of domination and of socially integrative behavior in teachers' contacts with children. *Child Dev.*, 1939, 10, 73-89.
- Anderson, H. H. and Brewer, Helen M. Studies of teachers' classroom personalities, I. Dominance and socially integrated behavior of kindergarten teachers. *Applied Psychol. Monogr.*, 1945, 6.



- Anderson, H. H. and Brewer, J. E. Studies of teachers' classroom personalities, II. Effects of teachers' dominative and integrative contacts on Children's classroom behavior. *Applied Psych. Monogr.*, 1946, 11.
- Anderson, H. H., Brewer, J. E., and Reed, M. F. Studies of teachers' classroom personalities, III. Follow-up studies of the effects of dominative and integrative contacts on children's behavior. *Applied Psych. Monogr.*, 1946, 11.
- Barr, A. S., et al. Second report of the committee on criteria of teacher effectiveness. *J. Educ. Res.*, 1963, 46, 641-658.
- Berlyne, D. E. *Conflict, arousal and curiosity*. New York: McGraw-Hill, 1960.
- Boardman, C. W. Analysis of pupil ratings of high school teachers. *Educ. Admin. Super.*, 1930, 16, 440-446.
- Book, W. F. The high school teacher from the pupils' point of view. *Pedag. Seminar*, 1905, 12, 239-288.
- Bowman, E. E. Pupil ratings of student teachers. *Educ. Admin. Super.*, 1934, 20, 141-147.
- Brookover, W. B. Person to person interaction between teachers and pupils and teaching effectiveness. *J. Educ. Res.*, 1940, 34, 272-287.
- Bruner, J. S. Personality dynamics and the process of perceiving. In R. R. Blake and G. V. Ramsey. (Eds.) *Perception: an approach to personality*. New York: Ronald, 1951, 121-147.
- Bruner, J. S., Goodnow, J. J. and Austin, G. A. *A study of teaching*. New York: Wiley and Sons, 1956.
- Bryan, R. G. Benefits reported by teachers who obtained written student reactions. *Educ. Admin. Super.*, 1942, 28, 69-75.
- Cogan, M. L. Theory and design of a study of teacher-pupil interaction. *Harv. Educ. Rev.*, 1956, 26, 315-342.
- Cotsonas, N. J. Jr., and Kaiser, H. F. A factor analysis of students' and administrators' ratings of clinical teachers in a medical school. 1962, *J. Ed. Psychol.*, 53, 219-223.
- Davenport, K. *An investigation into pupil rating of certain teaching practices*. Further Studies in Attitudes, Series VIII, Studies in Higher Education XLIX, Lafayette, Indiana: 1944.
- Davis, A. *Social-class influences upon learning*. Cambridge: Harvard University Press, 1948.
- Drawthorne, C. L. Relationship between pupil and student-teacher interaction and pupil ratings of teacher effectiveness. *Educ. Admin. Super.* 1954, 40, 283-286.
- Flanders, N. A. Analyzing teacher behavior. *Educ. Leadership*, 1961, b, 19, 173-180.
- Flander, N. A. Interaction analysis: a technique for quantifying teacher influence. *Research Mem.* BER-61-2, April, 1961a.
- Flanders, N. A. Personal-social anxiety as a factor in experimental learning situations. *J. Educ. Res.*, 1961, 45, 100-110.
- Flanders, N. A. Teacher influence in the classroom: I. Research on Classroom Climate. Unpublished manuscript, 1958.
- Flanders, N. A. Teacher-pupil contacts and mental hygiene. *J. Soc. Issues*, 1959, 15, 30-39.
- Foa, U. G. Convergences in the analysis of the structure of interpersonal behavior. *Psychol. Rev.*, 1961, 68, 351-353.
- Gage, N. L., Leavitt, G. S., and Stone, G. C. Teachers' understanding of their pupils and pupils' ratings of their teachers. *Psychol. Monogr.*, 1955, 69, No. 21, (whot No. 406).
- Gage, N. L. and Suci, G. Social perception and teacher-pupil relationships. *J. Educ. Psychol.*, 1951, 42, 144-152.



- Gold, H. A. The classroom isolate: an additional dimension for consideration in the evaluation of a quality education program. *J. Exp. Educ.*, 1962, 31, 77-80.
- Gombrich, E. M. On physiognomic perception. *Daedalus*. Winter, 1960, 228-241.
- Guthrie, E. R. The evaluation of teaching. *Training analysis and development of information*. Bulletin, 1963, 4, 199-206.
- Harman, H. H. *Modern factor analysis*. Chicago: University of Chicago Press, 1960.
- Jensen, L. E. An non-additive approach to the measurement of teacher effectiveness. In A. S. Barr et al. *Wisconsin studies of the measurement and prediction of teacher effectiveness*. *J. Exp. Ed.*, 1961, 30, 70-87.
- Keller, R. J. and Clark, K. E. *Student evaluation of classroom teaching*. Bureau of Ed. Res., Univ. of Minnesota, March, 1954, (mimeographed).
- Lawrence, P. R. *Organization, behavior, and administration*. Harewood, Illinois: The Dorsey Press and Richard D. Irwin, 1961.
- Lins, J. E. The prediction of teaching efficiency. *J. Exp. Ed.* 1946, 15, 2-60.
- McCall, W. A. *Measurement of teacher merit*. Raleigh, North Carolina: Superintendent of Public Instruction, No. 284, 1952.
- McClelland, D. C. *The achieving society*. New York: D. Van Nostrand Company, Inc., 1961.
- Morsh, J. E. and Wilder, Eleanor, W. *Identifying the effective instructor: a review of the quantitative studies, 1900-1952*. Chanute Air Force Base, Illinois: Training Aids Research Laboratory, Air Force Personnel, and Training Research Centre, Research Bulletin, October, 1954.
- Mowrer, O. H. *Learning theory and personality dynamics*. New York: The Ronald Press Company, 1960.
- Remmers, H. H. To what extent do grades influence student ratings of instructors? *J. Ed. Res.* 1930, 21, 314-317.
- Rokeach, M. *The open and closed mind*. New York: Basic Books, 1960.
- Root, A. R. Student ratings of teachers. *J. Higher Educ.*, 1931, 2, 311-315.
- Savage, Marjorie L. Changes in student teachers through the use of pupil ratings. Unpublished Ph.D. Dissertation, University of Illinois, 1959.
- Schmid, J. Factor analysis of the teaching complex. In A. S. Barr et al. *Wisconsin studies of the measurement and prediction of teacher effectiveness*. *J. Exp. Ed.* 1961, 30, 58-87.
- Steiner, I. D. Interpersonal behavior as influenced by accuracy of social perception. *Psychol. Rev.*, 1955, 62, 268-274.
- Thelen, H. A. Experimental research towards a theory of instruction. *J. Educ. Res.*, 1957, 45, 89-93.
- Thelen, H. A. Work-emotionality theory of the group as an organism. In S. Koch (Ed.) *Psychology: a study of a science*. Vol. III, *Formulations of the person and the social context*... New York: McGraw-Hill, 1959, 544-611.
- Ward, W. D., Remmers, H. H., and Schmalzried, N. T. The training of teaching personality by means of student ratings. *Sch. and Soc.*, 1941, 53, 189-192.
- Webb, W. B. and Nolan, C. Y. Student, supervisor, and self-ratings of instructional proficiency. *J. Educ. Psychol.* 1955, 46, 4-47.
- Wilson, W. R. Students rate teachers. *J. Higher Educ.*, 1932, 3, 75-82.

# ORGANIZATIONAL PLANS FOR TEACHING PHYSICAL EDUCATION IN ELEMENTARY SCHOOLS

S. BOYD TAYLOR AND WALTER H. WORTH

## Introduction

Traditionally, organization for staff utilization in the elementary school has centered around the idea of the self-contained classroom wherein one teacher instructs the pupils in all of the studies of a particular grade. Although widely used, it has been recognized that the self-contained classroom has certain weaknesses. Its use requires that each teacher must be capable of teaching all subjects, it tends to ignore the factors of aptitude and interest on the part of the teacher, and it makes no attempt to capitalize on special talents that may exist within a staff. Because of these weaknesses a number of alternatives to the self-contained classroom type of organizational plan have been proposed. These alternatives usually consist of some form of departmentalization, or the use of specialist teachers in subjects like art, music, and physical education.

In recent years, the centralization of rural school facilities and the rapid growth of population in most urban centers has meant that both rural and urban school systems are becoming larger and more complex. Increased size has, generally, meant larger school staffs, better facilities, and in some cases the availability of specialist services. In order to provide for the best possible type of instruction in large elementary schools, and in order to make effective use of facilities and available specialist and supervisory help, a re-examination of organization for instruction for the elementary school program as a whole, and for individual subjects, needs to be made.

Physical education is one subject which, due to the lack of qualifications or desire on the part of many teachers, may not always receive the attention it should under the self-contained classroom type of organization. In regard to organization for instruction in physical education, an elementary school may choose one of four general policies: (1) have each classroom teacher teach physical education to his own class, (2) make allowances for teachers to exchange classes so that those with special training and/or interest in physical education may relieve those who have not, (3) have physical education taught by a specialist who is a member of the school staff, and (3) have physical education taught by a visiting specialist employed by the unit or school system.

### Purpose

The study sought to accomplish the following five purposes: (1) to identify existing organizational provisions for teaching physical education in the elementary schools of Saskatchewan, (2) to determine the type of organizational plans preferred by elementary school teachers and administrators; (3) to determine the extent to which elementary school teachers feel competent in teaching physical education; (4) to compare the organizational plans used in large and small schools and in city and non-city school; and (5) to determine whether present practice, preferred practice, and teachers' feelings of competence are significantly related to the teacher variables of age, sex, and professional preparation in physical education.

### Procedure

#### *Collection of Data*

In order to gather information from principals and teachers for the study, two types of questionnaires, one for principals and one for teachers, were constructed and distributed to a randomly selected sample of thirty city schools and thirty non-city schools in the province of Saskatchewan. Only those schools having eight or more elementary classrooms in operation were included in the population studied. Responses were received from 57 principals and 588 teachers. This represented a 95 per cent return from principals and an 89 per cent return from teachers.

The teacher questionnaire asked for personal data about the respondents, such as age and sex, the grade level they were teaching, and the amount of professional preparation for teaching physical education. In order to determine the type of organizational plan presently being used for instruction in physical education, and the type of plan teachers would prefer for instruction in physical education, teachers were requested to check one item in each of the following sections of the questionnaire:

#### *I. Present Practice in Teaching Physical Education*

- (1) Teach physical education to my own class.
- (2) Physical education is being taught to my class by another elementary staff member.
- (3) Physical education is being taught by a special physical education instructor who is a member of the elementary school staff.
- (4) Physical education is being taught by a visiting specialist employed by the unit or school system.
- (5) No physical education is taught in my class.
- (6) Other arrangements. (Specify).

#### *II. Preference in Regard to Organization for Instruction in Physical Education*

- (1) Physical education taught by the classroom teacher.
- (2) Physical education taught by a classroom teacher in the school as part of a system of departmentalization or semi-departmentalization in which teachers "trade" subjects.
- (3) Physical education taught by a special physical education instructor who is a member of the school staff.



- (4) Physical education taught by a visiting specialist employed by the unit or school system.
- III. *Competence in Teaching Physical Education*
- (1) Feel proficient in teaching all areas of the physical education program.
- (2) Feel proficient in teaching most areas of the physical education program.
- (3) Feel inadequate in teaching most areas of the physical education program.
- (4) Feel inadequate in teaching all areas of the physical education program.

The principal questionnaire elicited basic data about the school. It asked principals to indicate whether any form of departmentalization or semi-departmentalization was being used for physical education instruction at different grade levels in the school. Principals were also requested to indicate the type of plan they would prefer for instruction in physical education at different grade levels in their school. The possible plans were the same as those listed in section II of the teacher questionnaire.

*Treatment of Data*

The sample studied was divided into sub-groups according to selected characteristics. The characteristics investigated were size of school, grade level that the respondents were teaching, and the teacher's age, sex, and amount of professional preparation for teaching physical education. Table I shows the size of the resulting sub-groups.

TABLE I

NUMBER OF RESPONDENTS CLASSIFIED ACCORDING TO  
SIZE OF SCHOOL, GRADE LEVEL TAUGHT, AGE, SEX AND  
AMOUNT OF PROFESSIONAL PREPARATION FOR TEACH-  
ING PHYSICAL EDUCATION

Characteristics	Principals N	Teachers N
Size of School		
Small (eight to twelve rooms) .....	40	334
Large (Thirteen or more rooms) .....	17	254
Grades Taught		
1-3 .....	.....	237
4-6 .....	.....	218
7 and 8 .....	.....	133
Age		
Under thirty .....	4	286
Thirty to forty-five .....	31	175
Over forty-five .....	22	127
Sex		
Female .....	5	455
Male .....	52	133
Professional Preparation in Physical Education		
No classes .....	16	181
One class .....	24	346
Two classes .....	12	47
Three or more classes .....	5	14

The significance of differences among the sub-groups in regard to present practice for teaching physical education, preferred plans, and expressed feelings of competence in teaching physical education were calculated by chi-square tests of independence. When the value of chi-square was found to be large enough to indicate a level of probability of .05 or lower ( $p \leq .05$ ), the difference was considered significant for the purposes of this study.

## Findings

### *Present Practice in Teaching Physical Education*

1. Approximately 79 per cent of the teachers in the sample schools were teaching physical education to their own classes. In almost all cases where teachers were not teaching physical education to their own classes it was being taught by another staff member as part of a system of departmentalization or semi-departmentalization. Less than 1 per cent of the classes were taught by a special physical education instructor who is a member of the school staff and in no instance was physical education being taught by a visiting specialist. In approximately 2 per cent of the classes there was no physical education being taught.

2. Principal's responses indicated that in 16 per cent of the sample schools there was some departmentalization being used for physical education instruction in grades one to three, in 53 per cent of the schools this form of organization was being used to some extent in grades four to six, and in 63 per cent of the schools it was being used to some extent in grades seven and eight. Although these percentages indicate that there is some departmentalization being used in many schools, teachers' responses showed that only 5 per cent of the teachers in grades one to three, 25 per cent in grades four to six, and 32 per cent in grades seven and eight were having their physical education taught by another staff member. These differences between grade levels were significant, indicating that there is more departmentalization or semi-departmentalization being used in the upper elementary grades.

3. The type of organizational plan being used was significantly related to both size and location of school. In larger schools there was a higher percentage of teachers who were using a plan whereby physical education was being taught by another staff member and there was a higher percentage of teachers using this type of organization in non-city schools than in city schools.

4. There was a significant relationship between the organizational plans used by teachers and the age of the teachers. There was a tendency for older teachers, in the schools studied, to have their physical education taught by another staff member.

5. There was no significant relationship between the type of organizational plans being used and the sex of the teachers.

6. Approximately 31 per cent of the responding teachers had taken no courses in physical education as part of their professional preparation, about 59 per cent had taken one course, 8 per cent had taken two courses, and slightly more than 2 per cent had taken three or more courses. There was a significant relation between the type of organizational plan being used by teachers and the number of physical education courses that teachers had taken. Teachers with more professional preparation for teaching physical education were more likely to be teaching physical education to their own classes, and possibly to classes other than their own.

7. Approximately 11 per cent of the teachers in the sample schools were teaching physical education to two or more classes. There was a significant relationship between the number of classes teachers were instructing in physical education and the teachers' age, sex, and amount of professional preparation in physical education. Of the teachers who were instructing two or more classes in physical education the majority were male, were under thirty years of age, and had taken two or more professional physical education courses.

#### *Organizational Plans Favored by Principals and Teachers*

1. The majority of principals indicated that they would prefer to have physical education taught by the classroom teacher in grades one to three, but would prefer a system of departmentalization or semi-departmentalization in grades four to six, and would like to see physical education taught to grades seven and eight by a special physical education instructor in the school.

2. The majority of teachers indicated that they would prefer to have physical education taught to grades one to three either by the classroom teacher or by a special physical education instructor in the school. For grades four to eight the majority of teachers preferred to have physical education taught by a special physical education instructor in the school.

3. There was a significant relationship between the sex of the teacher and the organizational plans favored for instruction in physical education. Although the majority of both male and female teachers favored an organizational plan whereby physical education could be taught by someone in the school other than the classroom teacher, a higher percentage of male teachers than female teachers preferred this type of organization, with a fairly large number of female teachers preferring to have physical education taught by the classroom teacher.



4. There was no significant relationship between the plans favored by teachers and the age of teachers.

5. There was no significant relationship between the plans favored by teachers and the amount of professional preparation that teachers had taken in physical education.

6. There was a significant relationship between the type of organizational plans preferred by teachers and the type of plan teachers were actually using. Although the majority of teachers preferred a plan whereby physical education is taught by a special physical education instructor in the school, teachers not choosing this type of organization tended to choose the type of plan that they were actually using at the time of the survey.

#### *Feelings of competence in Teaching Physical Education*

1. Approximately 8 per cent of the teachers in the sample schools indicated that they felt competent in teaching all areas of the physical education program; 53 per cent felt competent in teaching most areas of the physical education program; 34 per cent felt inadequate in teaching most areas of the program; and 5 per cent felt inadequate in teaching all aspects of the physical education program.

2. There was a significant relationship between the degree of competence expressed by teachers and the age of the teacher. Approximately 66 per cent of those under thirty years of age indicated that they felt competent in teaching most or all of the physical education program, while 57 per cent of the teachers in the thirty to forty-five age group, and 55 per cent of those who were over forty-five years of age indicated that they felt competent in teaching most or all of the physical education program.

3. There was no significant difference between male and female teachers in the degree of competence expressed towards the teaching of physical education.

4. There was a significant relationship between the degree of competence expressed by the responding teachers and the number of professional physical education courses that they had taken. The more professional preparation teachers had taken in physical education, the more competent they tended to feel in teaching physical education.

5. There was a significant relationship between the degree of competence expressed by the responding teachers and the type of plan that these teachers were using for instruction in physical education. Teachers who felt competent in teaching most or all of the physical education program were more likely to be teaching physical education to their own classes, while those who felt inadequate in teaching physical education were more likely to have

physical education taught to their classes by another staff member, or to have no physical education taught to their class.

6. There was a significant relationship between the degree of competence expressed by teachers and the type of plan that teachers preferred for instruction in physical education. Most of the teachers who preferred to have physical education taught by the classroom teacher were teachers who felt competent in teaching most or all of the physical education program, while only about one-half of those who preferred a type of organization in which physical education could be taught by another staff member, were teachers who felt competent in teaching most or all of the physical education program.

### Conclusions and Implications

An examination of the above findings gives rise to a number of conclusions and implications relative to physical education instruction in elementary schools.

#### *Organizational Provisions for Instruction in Physical Education*

This study indicates that the majority of principals and teachers in Saskatchewan's elementary schools would prefer a type of organizational arrangement for instruction in physical education other than the one which they are presently using. While it does not necessarily follow that existing arrangements are totally unsatisfactory, this does mean that many teachers are dissatisfied with present conditions and that both principals and teachers believe that physical education might best be taught using other organizational arrangements. Lacking objective evidence to indicate which type of organization is most suitable at different grade levels, the opinion of the majority of principals and teachers is probably sufficient reason to suggest that some changes should be made.

It was shown in the study, that approximately one-third of the teachers in elementary schools have had no professional preparation for teaching physical education and that over one-third of the teachers do not feel sufficiently competent in teaching physical education. It seems unrealistic to continue to expect these teachers to teach physical education adequately.

The study also indicates that in those schools where arrangements have been made for teachers to have their physical education taught by another staff member, the persons assuming responsibility for physical education tend to be better qualified and more competent than the teachers they are relieving. Having physical education taught by teachers who are better qualified and who feel competent in teaching this subject should result in students receiving better physical education instruction.



Both the opinion of principals and teachers as expressed in this study, and the literature on physical education, suggest that in the primary grades physical education should be taught by the classroom teacher. This study reveals, however, that most principals prefer a system of departmentalization or semi-departmentalization for instruction in physical education in grades four to six, while most teachers would prefer to have physical education taught by a special physical education instructor at this level, and that both principals and teachers prefer to have physical education taught to grades seven and eight by a special physical education instructor in the school. The relatively small size of many elementary schools and the lack of qualified physical education personnel means that, in most cases, it is unrealistic to hope to have physical education taught by a special instructor. It is possible, however to make provisions for teachers to trade subjects so that those with ability and interest in physical education may relieve other teachers of their responsibilities in this subject. It would seem most advisable then to have physical education taught by the classroom teacher in the primary grades and to make provision for, and encourage, teachers to trade subjects in grades four to eight so that physical education may be taught by those who are most capable of giving instruction in this subject. In large elementary schools, where it is feasible to do so, a teacher fully qualified to teach physical education could be employed to teach classes in grades four to eight and to give assistance in this subject to the teachers of grades one to three. A system of team teaching with one member of the team being highly qualified in physical education might also be used to advantage in the upper elementary grades.

In view of the fact that at the present time most teachers are expected to teach physical education to their own classes, and since this situation is likely to continue to exist for some time in many schools, every effort should be made to provide in-service help in physical education. This means that qualified personnel should be employed to act as supervisors or consultants in units and school systems. With the organization of larger administrative units and the consolidation of attendance areas this should be possible in most cases.

### *Teacher placement*

This study indicates that there are a considerable number of teachers in the schools who have taken two or more professional physical education courses. There are, in fact, enough of these teachers in Saskatchewan so that, theoretically, one could be placed in every school having eight or more elementary classrooms in operation. In addition to those teachers who have taken extra pro-



fessional preparation in physical education there are probably many teachers who because of background or interest, are very capable in teaching physical education. Through careful selection and placement of teachers on elementary school staffs it should be possible to have in each school at least one person who is capable of relieving teachers who cannot teach physical education satisfactorily.

It is shown in the study that younger teachers tend to feel more competent in teaching physical education than older teachers. This also has implications for teacher placement. By balancing school staffs in regard to the age of teachers, administrators will increase the possibility of having at least some teachers who are willing and capable in the field of physical education.

### *Teacher Education*

At the present time the majority of teachers in elementary classrooms are expected to teach physical education. As indicated by this study, many of these teachers have had no professional preparation for teaching this subject. As long as the present organizational arrangements in elementary schools exist, it will be essential that prospective elementary school teachers receive some preparation for teaching physical education. An increasing number of prospective elementary school teachers should also be encouraged to specialize in physical education so that they may assume physical education responsibilities in the upper grades of elementary schools. At the present time there are few people specializing in physical education and almost all of those who do become qualified in this subject are absorbed into high school staffs.

### *Further Research*

There is a need for research that would evaluate and compare the physical education programs offered under different types of instructional organization. The development of criteria to evaluate physical education programs offered in different schools would make it possible to relate effectiveness to organizational arrangements, as well as to additional factors such as facilities, teacher variables, supervisory assistance, and size and location of schools.

Similar studies should be carried out in other subject areas. The composite results of such studies might lead to the development of an organizational plan which would improve the effectiveness of the total educational enterprise at the elementary level.

# PREDICTIVE VALIDITY OF THE METROPOLITAN READINESS TESTS

HARVEY W. ZINGLE  
*Faculty of Education*

A. E. HOHOL  
*Superintendent of Jasper Place Public Schools*

## Introduction

In recent years the concept of readiness has come to occupy a prominent place in educational theory.

It is generally believed today that the assessing of a child's readiness for any aspect of learning is the first step that should be taken in successful teaching. The slogan "Begin where the child is" has gained great currency in educational practice and has, in many respects, revolutionized teaching methods and organizational plans in schools.

As the term readiness is used today it implies more than just emotional readiness, although that is important. It also includes physiological maturation, general mental ability, and experimental background or educational readiness. To be completely ready for an educational activity or learning experience, a child must want to learn, be sufficiently mature physiologically, possess appropriate mental abilities, and finally have had the right kind of educational experiences.

Thus, although most people agree that a child can profit sufficiently from attending school only when he has achieved minimum readiness, not all people agree (especially parents and school authorities) on an objective way of estimating readiness. Because of this inevitable conflict between parents and school authorities, most schools have in the past used the highly objective criterion of chronological age. The improvement of intelligence testing during the last couple of decades led most school systems to relax chronological age limits to permit the use of the mental age concept. Further sophistication in the field of tests and measurement has led many school systems to adopt a policy under which chronological age is combined with tests that attempt to measure the extent of the maturation of those skills considered most necessary for achievement in Grade I.

One such battery of tests is the Metropolitan Readiness Tests, which according to its authors(10) was devised to measure the traits and achievements of school beginners that contribute to their readiness for first-grade instruction.

### Problem

The publisher (10) of the Metropolitan Readiness Tests, after stating that the validity of their tests may be judged in terms of the correlation of "Readiness scores" with first-grade achievement, cite the results of a study done in Lebanon County, Pennsylvania, involving 487 cases. All of the 44 pupils rated as "Superior" on the Metropolitan Readiness Tests administered in September were found to be above the national norm in average "reading achievement," and all but one were above this norm in "Numbers" when tested in the following February on the appropriate tests of the Primary I Battery of the Metropolitan Achievement Tests. Of 46 pupils rated "Poor Risk" on the Metropolitan Readiness Tests, 22 were found to be below the national norm in average "reading achievement" and 39 below in "Numbers" when tested in February.

Gardner (Buros, 1963) who quarrels with the conclusions drawn from this study, notes that even if the norm is accepted as a minimal level of competence, which is certainly a debatable issue, about half of those who were considered "unready" for reading actually reached the norm of achievement. An unbiased interpretation of these data certainly would emphasize far more than does the publisher's manual, the extremely tentative meaning of the Metropolitan Readiness Test scores for individual diagnosis. This is especially crucial in the lower range where a difference in raw score of one may mean non-admission to school as opposed to admission.

At the same time educators have found that using chronological age or mental age is certainly no better. According to Dey (6) neither chronological age nor mental age are completely reliable as measures of the optimal age at which all children can profit from formal instruction. Bevington (1) in a study of 640 pupils in the Edmonton Public School System, including a large proportion of pupils under the age of six, found no correlation between chronological age and achievement.

It is dubious, then, whether the presently used methods of establishing whether or not a child is ready for formal instruction are adequate. At the same time the administrative decision regarding the child's admittance to school is usually irrevocable. Recognizing that such decisions have to be made, and further that in many schools these decisions are made on the basis of subjective judgments has led the authors of this study to undertake the testing of the following hypothesis:

1. The Metropolitan Readiness Tests are good predictors of success in Grade I reading and arithmetic, as measured by year-end standardized tests and year-end teacher ratings.



Two subsidiary hypotheses of this study were:

1. Teachers' year-end ratings tend to discriminate against boys more than standardized tests do.
2. Teachers' year-end ratings tend to discriminate against older students more than standardized tests do.

### Experimental Design

#### *Sample*

The subjects used in this study were all 545 of the Grade I students who were admitted to the Jasper Place Public Schools in the 1961-62 term. The first-grade readiness testing program of the Jasper Place Schools consists of the administration of all the Metropolitan Readiness Tests to all the candidates who seek first-grade admission. The legal school entry age in Jasper Place is six years, but any one who is at least five years and six months may write the tests. A total tests score of more than 40 on the Metropolitan Readiness Tests qualifies an under-age candidate for admission to Grade I. All children of legal age are admitted regardless of their test scores. Thus the age distribution of the Grade I students used in this study is negatively skewed. The actual distribution is shown in Table I.

TABLE I  
DISTRIBUTION OF SUBJECTS BY CHRONOLOGICAL AGE

Age .....	5-6	5-7	5-8	5-9	5-10	5-11	6-0	6-1	6-2	6-3	6-4	6-5	6-6	Over 6-6	Total
	19	23	38	32	48	46	49	52	48	41	45	31	27	46	545

#### *Procedure*

Even though the Metropolitan Readiness Tests are used as a screening device for only the underage candidates it was administered to all candidates in June 1961. In June 1962 the following tests were administered to all Grade I students:

1. The Marion Monroe Reading Test.
2. The Edmonton Public School Board Grade I Achievement in Arithmetic Test.

In addition to obtaining these test scores the teachers were asked to rate each of the students on a five point rating scale. The ratings were defined as indicated in Table II.

### Findings and Discussion

Intercorrelations of the predictor scores (Metropolitan Readiness scores) with the criterion scores (year-end achievement scores) for

the 545 subjects of this study are reported in Table III. Although the correlations are not as high as the correlations given in the test manual all are high and positive. To be fair to the test publisher, however, it must be pointed out that the children who made a total score below 40 were not admitted to school and so were not a part of the study. It is expected, therefore, that if these children had been retained the intercorrelations would have been somewhat higher.

TABLE II  
TEACHER RATING SCALE

Rating	Percentile Equivalent
1	90-99
2	75-89
3	25-74
4	10-24
5	1-9

TABLE III  
INTERCORRELATIONS OF METROPOLITAN READINESS  
TEST SCORES WITH YEAR-END MEASURES OF  
ACHIEVEMENT

Metropolitan Readiness Scores	Year-End Achievement Scores		
	Monroe Reading	Edmonton Arithmetic	Teacher Rating
Reading .....	.31	.45	.43
Number .....	.39	.53	.50
Total Score .....	.40	.56	.53

One rather interesting observation that can be made about the correlations in Table III is that the Metropolitan Reading Test predicts achievement in arithmetic (as measured by the Edmonton Arithmetic Test) significantly better than it does achievement in

reading (as measured by the Monroe Reading Test). This difference in correlations is significant at the .01 level of confidence.

Returning to the first hypothesis of this study, one would have to conclude that the Metropolitan Readiness Tests produce information which is useful when one attempts to make predictions regarding the academic success of Grade I students. In other words, the predictive validity of the Metropolitan Readiness Tests has been verified.

The first subsidiary hypothesis was stated as follows: Teachers' year-end ratings tend to discriminate against boys more than standardized tests do. This hypothesis was partially inspired by the fact that most studies (Samuels, 1943; Carrol, 1941; Donnelly, 1940) which show a sex difference in achievement at the Grade I level are based on teacher ratings. And one might suspect that the fact that first-grade teaching staffs are predominately female might be an ancillary factor of bias. It would appear, however, that since the point biserial correlation between sex and the Monroe Reading Test (— .11) and the point biserial correlation between sex and teacher rating (— .14) are both negative and significant this hypothesis is not upheld (Table IV). It would appear that there is a real sex difference in reading favoring the girls and also that this difference, which has been observed in the past, is not only a phenomenon of subjective teacher ratings but is also a phenomenon of standardized tests. It is also evident from an inspection of Table IV that this tendency to discriminate against boys is not a characteristic of the Edmonton Arithmetic Test ( $r = .00$ ).

TABLE IV  
POINT BISERIAL CORRELATIONS BETWEEN SEX AND  
YEAR-END MEASURES OF ACHIEVEMENT

	Year-End Achievement Scores		
	Monroe Reading	Edmonton Arithmetic	Teacher Rating
Sex .....	— .11	.00	— .14

The second subsidiary hypothesis was stated as follows: Teacher's year-end ratings tend to discriminate against older students more than standardized tests do. None of the correlations between age and year-end measures of achievement presented in Table V are significant. However, the correlation between age and Monroe Reading (0.6) differs significantly (.05 level of confidence) from the correlation between age and teacher rating (— .07). As the difference between the correlation of age with the Edmonton Arith-



metic Test is not significantly different from the correlation of age with teacher rating the second hypothesis is only partially supported.

TABLE V  
INTERCORRELATIONS OF CHRONOLOGICAL AGE WITH  
YEAR-END MEASURES OF ACHIEVEMENT

	Year-End Achievement Scores		
	Monroe Reading	Edmonton Arithmetic	Teacher Rating
Age .....	.06	.01	— .07

### Conclusions

1. The Metropolitan Readiness Tests are good predictors of success in Grade I reading and arithmetic as measured by year-end Standardized tests and year-end ratings. School administrators can no doubt find the Metropolitan Readiness Tests useful in helping them make admission decisions about pre-school children.
2. As both the Monroe Reading Test and the teachers' ratings tend to discriminate about equally against the boys it would appear that there is a real sex difference in reading.
3. The hypothesis that teachers' ratings tend to discriminate against older students more than standardized tests do was only partially upheld. The findings are very inconclusive as there seemed to be some indication of discrimination as far as reading is concerned but none in arithmetic.

### REFERENCES

1. Bevington, W., "Effect of Age of Entrance into Grade One on Subsequent Achievement." Unpublished Master's Thesis, University of Alberta, 1957.
2. Brodley, B. E., "An Experimental Study of the Readiness Approach to Reading." *Elementary School Journal*, 46, 1956, pp. 262-267.
3. Buros, O. K. (Ed.) *The Fourth Mental Measurements Yearbook*, Highland Park, New Jersey, The Gryphon Press, 1953.
4. Carrol, M. W., "Sex Differences in Reading Readiness." Unpublished Master's Thesis, Boston University, 1941.
5. Dean, C. D., "Predicting First Grade Achievement." *Elementary School Journal*, 39, 1939, pp. 609-616.
6. Dey, J. D., *Theory and Practice Governing the Time of School Entrance*. Monographs in Education No. 4, The Alberta Advisory Committee on Educational Research, University of Alberta, Edmonton, 1960.
7. Donnelly, H. E., "A Study of Word Recognition Skills in Grade One." Unpublished Master's Thesis, Boston University, 1940.
8. Gates, A. I., "The Necessary Mental Age for Beginning Reading." *Elementary School Journal*, 37, 1937, pp. 497-508.
9. Hildreth, G., *Readiness for School Beginners*. Yonkers-on-Hudson, New York, World Book Co., 1950.
10. *Metropolitan Readiness Tests*, Yonkers-on-Hudson, New York, World Book Company, 1949.
11. Samuels, F. L., "Sex Differences in Reading Achievement." *Journal of Educational Research*, 36, 1943, pp. 594-603.

# RELATION OF TEXTBOOK DIFFICULTY TO MATHEMATICS ACHIEVEMENT IN JUNIOR HIGH SCHOOL

L. DOYAL NELSON

The importance of the role of the textbook in determining the content, organization, and mode of presentation of school mathematics courses can hardly be over-estimated. The production of textbooks in recent years with content strikingly different from that of traditional textbooks has been largely responsible for the current revolution in school mathematics. This change in content has brought with it the question of how the material can best be organized and presented to the students. In other words, is there any one pattern of organization or method of presentation of mathematics ideas through textbooks which will foster greater learning efficiency on the part of the pupil?

With our present knowledge of the nature of the learning process it is doubtful if we will find a definite answer to this question which would apply to all pupils. However, if there were two sets of mathematics textbooks which contained essentially the same material and which differed only in organization and method of presentation, it would be possible to shed some light on the question for pupils of certain levels of ability. This could be done by taking two groups of pupils of comparable ability, assigning one set of textbooks to one group of pupils, and the other set to the other group, and then comparing the mathematics achievement of these two groups of pupils after they had used the texts for a specified period of time.

One of the most influential organizations in the current revolution in school mathematics is the School Mathematics Study Group (SMSG). They have focussed their attention on the production of textbooks with content in line with the needs of pupils in a modern technological society. They have produced for both the grade seven and grade nine levels two sets of textbooks in which the content is similar but which differ in reading difficulty, organization, and method of presentation. The text *"Mathematics for Junior High School, Volume 1"* was prepared especially for college-bound seventh grade students. A modification in the presentation and organization of much of the material in this text was subsequently made for slower learning seventh grade pupils and is contained in the text called *"Introduction to Secondary Mathematics, Volume 1"*. According to an SMSG newsletter "The changes and adjustments which

were made were prompted by a desire to simplify the presentation and reduce the reading difficulty. Explanatory sections were shortened and exercises added to lead pupils through simple steps to appropriate conclusions<sup>1</sup>". These texts contain a modern treatment of such topics as systems of notation, the system of whole numbers, rational numbers, factors and primes, non-metric geometry, and so forth.

In a similar manner the textbook "*Mathematics for High School, First Course in Algebra*" was prepared for college-bound ninth grade pupils and the organization and presentation of its content was subsequently modified for slower learning pupils. The modified version appears in the textbook called "*Introduction to Algebra*". These texts contain a modern treatment of such topics as: the system of real numbers, properties of operations and order, sets and sentences, polynomial and rational expressions, functions and the like. For a complete picture of the nature of the content and nature of the modifications the reader should refer to the textbooks themselves.<sup>2</sup>

### *Problem*

It was the object of this investigation to study the following questions: What would be the effect on the mathematics achievement of high ability students who used the SMSG texts designed for slower learning pupils? Would their mathematics achievement be different from that of similar high ability students who used the SMSG sets designed for college-bound capable students?

Specifically, the basic hypotheses to be tested were as follows:

(a) There is no difference in the mathematics achievement, as measured by a standardized test, between high ability grade seven students who use the seventh grade SMSG text for college-capable students and those who used the text for slower learners.

(b) There is no difference in the mathematics achievement, as measured by various unit tests prepared specifically to test achievement in material covered by the texts, between high ability grade seven students who use the seventh grade SMSG text for college-capable students and those who use the text for slower learners.

Similar hypotheses were tested for high ability ninth grade pupils.

### *Design*

This investigation was conducted in fourteen schools in Minnesota during the 1961-62 school term. Each of the schools provided two mathematics classes—both of which consisted of either high ability seventh grade pupils or high ability ninth grade pupils.

<sup>1</sup> Newsletter No. 11, School Mathematics Study Group, Leland Stanford University, 1962; page 15.

<sup>2</sup> These textbooks may be obtained from: Yale University Press, School Mathematics Study Group, 92A Yale Station, New Haven, Connecticut.



One of the classes in each school used the SMSG text for college-capable students at the appropriate grade level and the other class in the school used the SMSG text which had been modified for slower learners.

Seventh grade pupils for the two experimental classes in each school were selected from among those grade seven students in the school who were above the mean in measured mental ability. Ninth grade pupils for the experiment were selected from among the top third in mental ability. Teachers and administrators in each school were asked to use official files and their knowledge of pupil ability to make up the two classes so that their mean abilities were as nearly alike as possible. Six pairs of seventh grade classes and eight pairs of ninth grade classes were obtained for the experiment.

One mathematics teacher from each school was assigned to teach both experimental mathematics classes. Teachers were instructed to use a method of presentation for each class which would conform closely to the textbook assigned. Every attempt was made to keep other variables as nearly constant as possible. For example, teachers for the experiment were chosen by the Minnesota National Laboratory and throughout the experimental period maintained close contact with members of this organization. Officials of the Laboratory met with the teachers of the experimental classes once each month, helped them with instructional problems, and consistently stressed the necessity of controlling such variables as time spent in developing ideas, time spent in review, standards in scoring tests, use of motivating devices, and so forth. The teachers involved were not typical of junior high school mathematics teachers. All of them had had previous experience teaching experimental programs, and all but one of them had taken formal classes specifically designed to help them to interpret modern programs in school mathematics. The total number of quarter credits they held in college mathematics varied from 24 to 67.

There was nothing in the design of this experiment to guarantee that a student who was classified as high ability in one school would be so classified in another. This lack of cross classification made it necessary to consider each pair of classes in each school to be a separate experimental situation for the purpose of testing the hypotheses. In the six pairs of seventh grade classes there were 285 students and in the eight pairs of ninth grade classes there were 460 students.

Various tests were administered during the course of the experiment and the resulting measures provided the basis for testing the hypotheses which have already been stated. In order to get measures of pre-experiment mathematics achievement, STEP Mathe-

matics, Form 3B was administered to all seventh grade participants and STEP Mathematics, Form 2B to all ninth grade participants during September, 1961. A series of five unit tests designed to measure achievement in topics specifically treated in the SMSG textbooks was administered to pupils at each level during the course of the year, as the topics were completed. The five unit tests administered to seventh grade students covered the following topics: non-metric geometry, factors and primes; rational numbers and fractions; decimals, ratio, and percent; measurement, area and volume; parallels, polygons, prisms, circle statistics and graphs. The five at the ninth grade level covered the following topics: sets, sentences, and variables; open sentences and properties of operations; real numbers, multiplication and addition of real numbers; properties of order, subtraction and division of real numbers; factors, exponents, radicals and polynomials.

At the termination of the experimental period in May, 1962, STEP Mathematics, Form 3A was administered to seventh grade students and STEP Mathematics, Form 2A to all ninth grade students. The unit tests scores and the scores on the final STEP Mathematics Tests were used as criteria of mathematics achievement.

### *Statistical Treatment*

A combination of analysis of covariance and regression analysis was used to treat the data collected. The covariable at each grade level was the score on the pre-experiment STEP Mathematics test. It was thus possible to take into account any pre-experiment differences in mathematics achievement between the groups being compared. Criterion measures were scores on the post-experiment STEP Mathematics test and on the five unit tests administered at each grade level. As explained before, each pair of classes in each school was considered to make up a separate experimental situation or replication of the experiment. Thus there were thirty-six tests of hypotheses at the grade seven level: one for each of the six pairs of classes when the set of final STEP Mathematics scores was used as a criterion and five for each of the six pairs of classes when each of the unit test scores was used as a criterion. At the grade nine level there were forty-eight tests of the hypotheses; one for each of the eight pairs of classes when the set of final STEP Mathematics scores was used as a criterion and five for each of the eight classes when each set of the unit test scores was used as a criterion.

It was assumed that an individual student's score on any of the criterion tests was a function of his pre-experiment achievement and the particular SMSG textbook he used during the experimental period.



The model for this condition may be expressed as follows:

$$Y_a = B_a X_a + G + I_a$$

$$Y_b = B_b X_b + G + I_b$$

where:

$Y_a$ : the criterion score of any student using the SMSG mathematics text for the college-capable

$Y_b$ : the criterion scores of any student in the same school using the modified SMSG mathematics text

$B_a X_a$  and

: the regression effect of the covariable

$B_b X_b$

$I_a$ : the effect of using the SMSG mathematics text for college-capable students

$I_b$ : the effect of using the modified SMSG mathematics text

$G$ : error components which were assumed to be normally distributed with mean zero and variance equivalent to that of the populations from which the samples were drawn.

Normally, when analysis of covariance is used the means of the groups being compared are adjusted and an appropriate test used to determine whether the adjusted means are significantly different. However, this procedure can be used only when the slopes of the two regression lines involved are not different. In this investigation it was important that achievement as measured by the criterion tests could be studied for the entire range of the covariable. Thus the hypothesis tested was as follows:

$$H_0: \begin{cases} B_a = B_b \\ I_a = I_b^3 \end{cases}$$

An appropriate test of this hypothesis could be used to determine simultaneously whether the slopes of the regression lines and the y-intercepts were different. These tests were based upon the calculation of an  $F$  value obtained as follows:

$$F = \frac{\frac{s_r^2 - s_a^2}{n_r - n_a}}{\frac{s_a^2}{n_a}}$$

where:

$s_r^2$ : is the sum of squares of deviations of each observed score on the criterion for both classes being compared from a regression line computed from the combined data for the two classes.

<sup>3</sup> For a complete discussion of the procedure see: Henry Scheffe—*The Analysis of Variance*, John Wiley & Sons Inc., 1958 (esp. pp. 68-72).



- $s_a^2$ : the sum of square of deviations of observed criterion scores for the class using the SMSG text for the college-capable combined with the sum of squares of deviations of observed criterion scores for the class using the modified SMSG text. Each set of deviations was taken from the regression line of a particular criterion variable on the initial STEP scores for each class taken individually.
- $n_a$ : the sum of degrees of freedom for class using the SMSG text for college-capable students and the degrees of freedom for the class using the modified SMSG text.
- $n_r$ : number of degrees of freedom available if data for the two classes being compared is combined.

Wherever the hypothesis of no difference was rejected, scatter diagrams of the data involved were drawn to permit a complete analysis of the nature of the differences.

Results

Seventh Grade

Table I provides a summary of the results of the various tests of the hypothesis for the grade seven classes. The level for rejecting the hypotheses was set at 5 per cent.

TABLE I  
SUMMARY OF RESULTS OF TESTS OF HYPOTHESES FOR SEVENTH GRADE CLASSES

SCHOOL	CRITERION					STEP 3A
	Unit Test 1	Unit Test 2	Unit Test 3	Unit Test 4	Unit Test 5	
1 .....	a	a	a	a	a	a
2 .....	r	a	a	a	a	a
3 .....	a	a	a	a	a	a
4 .....	a	a	a	a	a	a
5 .....	a	a	a	a	a	a
6 .....	r	r	r	a	r	r

a—hypothesis accepted  
r—hypothesis rejected

It will be noted that five of the six rejections occurred in School 6. The scatter diagrams for these five rejections indicated that there was a tendency for the high ability students in this school who were low achievers in mathematics at the beginning of the experiment to achieve better if they used the modified SMSG mathematics text rather than the one for the college-capable students. There was also a tendency—though not so marked—for those who had been high achievers at the beginning of the experiment to achieve better on the criterion tests if they used the SMSG text for college-capable students than if they used the modified

version. This same trend was also noted for Unit Test 4 in this school though the differences were not great enough to be statistically significant. Of the remaining 29 of the cases in which the hypothesis was accepted, 14 of them showed a trend similar to that described above. In the case of Unit Test 1, for School 2 where the hypothesis of no difference was rejected, the pupils who used the modified version of the SMSG grade seven text tended to achieve better, over the entire range of the covariable, than those who used the regular text.

*Ninth Grade*

The results of the various tests of the hypothesis at the ninth grade level are presented in Table II. Here again the level for rejecting the hypothesis was set at 5 per cent.

TABLE II  
SUMMARY OF RESULTS OF TESTS OF HYPOTHESES FOR  
NINTH GRADE CLASSES

SCHOOL	CRITERION					
	Unit Test 1	Unit Test 2	Unit Test 3	Unit Test 4	Unit Test 4	STEP 2A
7 .....	a	r	a	a	a	a
8 .....	a	r	r	r	a	a
9 .....	a	a	a	a	a	a
10 .....	a	r	a	a	a	a
11 .....	a	a	a	r	r	a
12 .....	a	a	a	a	a	a
13 .....	a	a	r	a	r	a
14 .....	a	r	a	r	a	a

a—hypothesis accepted  
r—hypothesis rejected

When the STEP mathematics post-experiment scores were used as criterion there was no significant difference in any of the schools between the achievement of pupils using the SMSG text for the college-capable and those using the modified version of this text. Of the forty remaining tests of the hypothesis when the unit test scores were used as criteria twenty-nine were accepted and eleven were rejected. For School 7 and Unit Test 2, regression analysis revealed that students who used the modified version of the SMSG tended to achieve better over the whole range of the covariable than students who used the SMSG text for the college-capable. In School 8 there were three rejections on the basis of unit tests and in all these cases the tendency was for students who used the modified SMSG text to score higher than the students who used the SMSG text for the college capable. This tendency was most marked for those students whose pre-experiment mathematics achievement was high. For the seven remaining rejections, the same pattern as noted

for seventh grade pupils was found. That is, the lower achievers tended to score higher on the unit tests if they used the text for college-capable students. In most cases this tendency decreased for the higher achieving high ability students. However, even for the highest achieving students, using the text for the college-capable did not appear to give them any decided advantage. At this grade level in the majority of cases where the hypotheses of no difference were accepted the trend—though not great enough to be statistically significant—was for low achievers to perform better in the criterion tests if they used the modified SMSG text. This tendency, in general, decreased for higher achieving pupils.

### *Conclusions and Recommendations*

1. The number of times the hypothesis that there was no difference between the mathematics achievement of high ability pupils who used the SMSG text for slower learners was rejected indicated an effect that could not be entirely attributed to chance.

2. In the cases where the hypothesis was rejected the most common tendency was for the lower achievers among these high ability pupils to achieve better on the criterion tests if they had used the SMSG text for slower learners. The superiority of the modified text tended in many cases to decrease for the higher achieving, high ability pupils and at the grade seven level there was some evidence to indicate that using the modified text would actually have an adverse effect upon the subsequent achievement of the most capable of these high ability pupils.

3. The variability in the nature of the differences as revealed by regression analysis would indicate that there are factors other than level of textbook difficulty which affect the mathematics achievement of high ability pupils at these levels. The factors are undoubtedly associated with the teacher, the pupils, the school, or various combinations of these.

4. Since the modified versions of the SMSG texts at both the grade seven and grade nine levels tended to facilitate the learning of mathematics on the part of high ability, lower achieving pupils involved in this experiment, there can be little doubt that modifications of the kind made by the School Mathematics Study Group would assist lower ability students in learning mathematics.

5. No attempt was made in this investigation to determine the relative effectiveness of the various devices used in the modified SMSG texts to simplify presentation. There is a need for this kind of study. There is also a need to determine which devices are most effective for simplifying the presentation of specific topics.

6. The term "high ability" in this investigation was not rigorously defined. There is a need for more information about the com-



parative performance of students within specifically-stated ranges of ability when they use the SMSG mathematics textbook for college-capable students or the modified version designed for slower learners. There is a need for more information about the comparative performance of students of various levels of ability on specific topics in mathematics presented in various ways. Indeed, one might question the need for providing a more difficult version of a mathematics textbook for college-capable students. Among the high ability students in this investigation only the highest achievers seemed to find the treatment which involved the use of the SMSG text for the college-capable advantageous from the standpoint of mathematics achievement. This advantage was slight and the number of students involved was small. It might be argued that, at the junior high school level, textbook writers should continually search for better ways of making the presentation of mathematical ideas as simple as possible. It would appear, for example, that reading difficulties should not be allowed to interfere with the acquisition of fundamental ideas in mathematics, that examples should be provided to promote pupil discovery of mathematical relationships and principles; and that a variety of significant problem situations should be provided to enable the student to make applications of the relationship and principles.

# THE HUMAN FACTOR IN CRITICAL ANALYSIS

NORMAN CUTHBERTSON

## I. Introduction

A prime characteristic of teaching has been its resistance to change. Most teachers, it seems, teach as they were taught. Moreover, teaching has been regarded as a rather simple task involving little more than the imparting of knowledge in neatly packaged quantities to usually receptive children. The assumption being that, if the teacher knew his subject, little else was required. Good teaching, therefore, was easily defined. Its hallmarks were faithful adherence to a prescribed curriculum and a devotion to exposition as a teaching technique. Schwab (7) describes the traditional dogmatic classroom as one in which "the role of the teacher was to explain what the book left unclear and to test the student's grasp of what he was told."

Behind the close door of her classroom the teacher taught as she chose or, as was more often the case, as best she knew how. The excitement of the sixties, in education, centers around an effort to break through these dual concepts of insularity and self-sufficiency which long have characterized the role of the teacher in America education. Such a breakthrough, it is hoped, would help to effect an improvement in the learning situation for pupils. Team teaching seeks to achieve this end by giving groups of teachers joint responsibility for the education of a large number of students. Under such a means the onus for improvement rests on the group, in that it assumes the prime responsibility for meeting the needs of all students under its care. Critical analysis, on the other hand, seeks to batter down the walls of complacent insularity which have long encompassed teaching to lay bare the facts of the matter. Its technique of analysis, based on data gathered by direct observation of the teacher in action, tends to focus directly on the individual concerned. Solely accountable for the amount of learning that occurs in his classes, the teacher is placed under much greater stress than is the case when members of a group are held jointly responsible.

## II. Delineation and Potential

Let us now turn to a brief delineation of the procedure previously referred to as critical analysis and to some consideration of its potential. In the first place, it is not intended as a one-shot treatment. Rather it should be regarded as a continuing process which should not be instituted until a climate conducive to growth

is established. The pattern proposed at the Harvard-Lexington Summer Program includes a six-step supervisory cycle consisting of: (1) planning a lesson with a teacher, (2) observation of the teaching, (3) analysis of the teaching, (4) critique or analysis session with the teacher observed, (5) observation follow-up the next day, and (6) critique follow-up the next day. Ideally the cycle would be repeated as often as it was deemed necessary or as it was productive of the desired results. Planning of the lesson with the teacher makes it possible for the supervisor to become familiar with the objectives of the lesson as well as with the content and the strategies to be employed. In the second phase it is intended that the supervisor remain in the classroom activity observing the teacher throughout the progress of the lesson. Thereafter, time should be provided for both the teacher and the supervisor to collect their thoughts but, in particular, for the supervisor to analyse the exhaustive notes or transcriptions taken in an attempt to discover patterns of teacher behavior that bore directly on the success or failure of the lesson. The strategies employed by the teacher to achieve his objectives must be examined in the same light. In the ensuing face to face critique or analysis session, the supervisor attempts to present the evidence gathered in a manner conducive to self-analysis on the part of the teacher. Thus, it is hoped that by making the teacher aware of his strengths and, more particularly, of his weaknesses in the highly charged atmosphere of the analysis session, he will be strongly motivated to improve his methods. But the reactions of teachers to this kind of stress are as yet undertermined and the effects of continued exposure unknown. Be that as it may, the need to update and upgrade teaching is so great that we must be prepared to face the dangers. Indeed it may not be always possible or desirable to separate instruction from instructor, and the strength of the procedure may lie in the acceptance of this fact by both parties. If we accept Alpert's(1) suggestion that dominant motives in a motivational hierarchy are aversive or tensional in nature, we need not be too concerned about any momentarily adverse effects a discreetly conducted analysis session may produce. However, since the supervisor must assume responsibility for the outcome, he must be aware of the dangers inherent in the technique.

This paper, therefore, proposes to examine the reactions of teachers at the 1963 Harvard-Lexington Summer Program to their initial exposures to critical analysis in an effort to test the hypothesis that instruction is not likely to improve unless it is subjected to continued analysis.



III. Presentation and Analysis of Data

A research specialist examining this paper and discerning a curious combination of the techniques of direct observation and use of a questionnaire for gathering the data here presented will, it is hoped, recognize the need to gather as much data as possible in the few days available to the writer. So too, will he recognize that the author has drawn heavily on his own experience, both past and present, in arriving at conclusions.

Tables I and II present in summary form the responses of teachers to the first set of items in each of parts one and two of the questionnaire.

TABLE I

In the following check the item that best describes how you felt on the first occasion that your teaching was subjected to critical analysis at the 1963 HLSP.

Extremely nervous .....	11%	} 22%
Very nervous .....	11%	
Somewhat nervous .....	29%	} 50%
A little nervous .....	42%	
Not at all nervous .....	8%	

TABLE II

In the following check the item that best describes how you now feel when your teaching is being subjected to critical analysis. (At end of second round of teaching.)

Extremely nervous .....	0%	} 8%
Very nervous .....	8%	
Somewhat nervous .....	18%	} 74%
A little nervous .....	42%	
Not at all nervous .....	32%	

Note from Table I than 22% of the responses lie in the upper two categories which indicate a highly nervous state on the part of the teacher whose work is being subjected to analysis. On the other hand, 50% of the responses lie in the lower two categories indicating little or no nervousness. From this we might conclude that the suggestion relative to the effects of stress made earlier in this paper is not borne out. Therefore, let us consider Table II which indicates a strong shift in favor of little or no nervousness. The corresponding figures are 8% for the upper two categories and 74% for the lower two. This shift is statistically significant. Thus, it appears that the feeling of nervouness experienced by teachers at Harvard-Lexington on their first exposure to critical analysis diminished rapidly as the technique became more familiar. It might also be pointed out that none of the respondents indicated extreme nervousness after his second or third experience with critical analysis. Even greater significance may be attached to this shift when consideration is given to the fact that the questionnaire was distributed from one to three weeks after the respondents

had their first analysis sessions, and the time lapse may have contributed to some modification in their recollection of the event.

One further item is worthy of record. Teachers have frequently commented that they were more nervous during the observation of their work than during the critique. Three respondents took the time to indicate this on the side of the questionnaire. Here, then, is an area that may prove fruitful for further research.

Tables III, IV, V and VI are grouped to make it possible to observe whether or not the trend indicated in Tables I and II is continued. Although the items from which the data are assembled were included in the questionnaire primarily as reinforcement for the first ones, they have some individual merit.

TABLE III

In the following check the item that best describes how you felt on the first occasion that your teaching was subjected to critical analysis at the 1963 HLSP.

Extremely apprehensive .....	11%	} 18%
Very apprehensive .....	7%	
Somewhat apprehensive .....	29%	} 53%
A little apprehensive .....	49%	
Not at all apprehensive .....	4	

TABLE IV

In the following check the item that best describes how you now feel when your teaching is being subjected to critical analysis (at end of second round of teaching).

Extremely apprehensive .....	0%	} 11%
Very apprehensive .....	11%	
Somewhat apprehensive .....	22%	} 67%
A little apprehensive .....	32%	
Not at all apprehensive .....	35%	

TABLE V

In the following check the item that best describes how you felt on the first occasion that your teaching was subjected to critical analysis at the 1963 HLSP.

Extremely insecure .....	11%	} 25%
Very insecure .....	14%	
Somewhat insecure .....	11%	} 65%
A little insecure .....	36%	
Not at all insecure .....	29%	

TABLE VI

In the following check the item that best describes how you felt when you teaching was being subjected to critical analysis (at end of second round of teaching).

Extremely insecure .....	0%	} 4%
Very insecure .....	4%	
Somewhat insecure .....	25%	} 71%
A little insecure .....	25%	
Not at all insecure .....	46%	

Excluding minor variations, the trend indicated in the first two tables is again evident. Since the median response in Tables I, III and V lies between the "somewhat" and "a little" categories, it appears that the teachers generally were not unduly nervous, apprehensive or insecure on the first occasion that they experienced critical analysis. Moreover, the shift in this median in Tables II, IV and VI to the middle of the "a little" category is statistically significant. If so favorable a shift in teacher reactions occurred in from one to three weeks, it may be assumed that a long-term program of analysis aimed at the improvement of instruction might early reach the stage where these kinds of emotional reactions are not longer a major consideration.

The shift in the two upper levels of from 25% in Table V to 4% in the Table VI is most significant. It appears that much of the insecurity felt by teachers during the initial sessions was due to uncertainty about what was going to happen. Once the teacher had experienced analysis, and had become aware of its purpose, feelings of insecurity faded quickly. This is further born out by the fact that 46% of all respondents indicated that they felt "not at all insecure" in later sessions.

Comparing Table IV with Tables II and VI, we note that the range from 11% to 67% for the spread from "extremely" and "very apprehensive" to "little" and "not at all apprehensive" is not so great as it is for the same categories dealing with feelings of nervousness and insecurity. These two range from 8% to 74% and from 4% to 71%, respectively. The question which now arises relates to the significance we can attach to this difference. A re-examination of the questionnaires reveals that at least three of the respondents felt more apprehensive during their later experiences with critical analysis, but none felt more nervous or more insecure. Two reasons for this increased feeling of apprehension appear valid. The first is that the teacher had not anticipated the probing nature of the analysis session and was apprehensive about further revelations of weakness in his teaching. The second, and more plausible one, is that the teacher reacted unfavorably to the strategy used by the analyst and was apprehensive about the conduct of future sessions.

The evidence to date appears to negate the contention that the danger of unfavorable reaction among teachers would tend to severely restrict the use of critical analysis. On the contrary, we have evidence that most teachers at the 1963 Harvard-Lexington Summer Program adjusted readily to his new approach to the problem of the improvement of instruction.

Let us now turn to an examination of the data provided by the second half of the questionnaire.



TABLE VII

Indicate the extent that your teaching has changed while you have been at HLSP 1963 by checking the appropriate item.

No change .....	7%	} 36%
A little change .....	29%	
Some change .....	25%	} 39%
A great amount of change .....	25%	
An extreme change .....	14%	

Since 93% of all teacher polled indicated that their teaching had changed, and since the median response fell in the middle or "some change" category, it may safely be assumed that there were factors at work at Harvard-Lexington which contributed to this change. The next item on the questionnaire was included in an effort to assess the importance of the factor of critical analysis to this process of change. The teachers were asked to rank the following seven items in order of their influence on any change that may have occurred:

- 1. Discipline meetings with subject specialists.
- 2. Cross discipline team meetings.
- 3. Lectures at Estabrook.
- 4. Observation of analysis sessions of other teachers.
- 5. Critical analysis sessions of your teaching.
- 6. Observation of other teachers in action.
- 7. Informal discussions with other teachers.

Table VIII presents the tabulation of responses with respect to the influence of critical analysis only.

It will be noted that 68% of all respondents ranked critical analysis of their teaching as being among the first three influences on any changes that may have occurred in their teaching while they were students at Harvard-Lexington. Conversely only 20% ranked it lower than fourth in influence. Although critical analysis was ranked well ahead of all other influences, it is worthy of note that the opportunity to observe other teachers in action was ranked second in importance.

TABLE VIII  
Rank order for critical analysis.

First .....	25%	} 68%
Second .....	18%	
Third .....	25%	
Fourth .....	11%	} 20%
Fifth .....	14%	
Sixth .....	3%	
Seventh .....	3%	

At the end of the questionnaire, space was reserved for any comments teachers cared to make about their personal reaction to the type of analysis experienced at Harvard-Lexington. Only 25% of the respondents chose not to comment! Of the 25% whose comments could be classed as unfavorable, about two-fifths felt that the supervisors did not show sufficient concern for the feelings of

teachers. Another two-fifths felt that the sessions were superficial and did not probe deeply enough. About 35% of all comments were favorable, with the remaining 15% vacillating between a modicum of praise and a neutral position. Thus, the statement of the observation team leader who commented that 40% success was a rather good batting average appears to be a valid assessment. However, in our jubilation over our successes we must not overlook the distress of the minority who chose to express concern over the apparent lack of consideration for the feelings of teachers.

Finally, from Table IX, we can obtain some conception of who the person might be whom teachers would be most favorably disposed to accept as a critic of their teaching.

TABLE IX

If critical analysis were to become a regular feature of the supervisory program in your school system this fall, whom would you prefer as your supervisor Check one.	
29%	the principal of your school.
29%	a departmental head or subject specialist from your school.
3%	a team leader in your school.
14%	a supervisor employed by the school system.
11%	an independent person from outside the system.
14%	other .....

By combining the first three items we note that 61% of all respondents prefer to be supervised by someone attached to the staff of the school. Add to this the responses under "other" which favored a team of the teacher's peers and the figure reaches the two-thirds point. The small vote of only 3% for a team leader may be attributed to lack of experience, on the part of most respondents, with team teaching. Similiary, the light vote of 11% for an independent person indicates little interest due to lack of experience with this type of supervision. How, then, may one interpret the meager 14% response in favor of a supervisor employed by the school system? Certainly this adds credence to the notion that teachers generally have an unfavorable image of supervisors. Furthermore, four respondents added some such comment as "it depends on the person." It would appear, therefore, that the selection of the person to be charged with the responsibility of using the technique of critical analysis will be a most important factor in the success of a program.

IV. Summary and Conclusions

It has been frequently intimated that the writer was agreeably surprised by the fact that teachers generally did not react unfavorably to critical analysis. Certainly the evidence is sufficiently encouraging to merit continued experimentation and, as trained personnel become available, to begin its use in a more direct attack

on the problem of the improvement of instruction. But there are other considerations which must be given high priority if the technique of critical analysis is to reach its potential. First among these is the need to upgrade supervision itself in the hope of improving the image of the supervisor in the eyes of teachers. A second consideration is that of bringing teachers to an acceptance of supervision as a professional necessity. Teachers, it appears, want to improve their teaching. Therefore, it behooves the supervisor to improve his professional competencies to the point where his willingness to work with a teacher in an honest effort to seek improvement in the level of instruction is acceptable to the teacher. This approach is quite compatible with the dualism of technical and human skills which Lucio and McNeil(4) believe to be essential for intelligent activity on the part of the supervisor.

## REFERENCES

1. Alpert, Richard, *The Shaping of Motives For Learning*. (Mimeographed).
2. Bartky, John A., *Supervision As Human Relations*. Boston: D. C. Heath and Company, 1953.
3. Bruner, Jerome, *The Process of Education*. Cambridge: Harvard University Press, 1960.
4. Lucio, W. H. and McNeil, J. D., *Supervision—A Synthesis of Thought and Action*. McGraw Book Company, Inc., 1962.
5. Morse, Arthur D., *Schools of Today, Tomorrow*. Garden City, New York: Doubleday and Company, Inc., 1960.
6. Reeves, Arthur and Melsness, Harold and Cheal, John, *Educational Administration—The Role of the Teacher*. Toronto: The MacMillan Company of Canada, Ltd., 1962.
7. Schwab, Joseph and Brandwein, Paul, *The Teaching of Science*. Cambridge: Harvard University Press, 1962.
8. Selltz-Jahoda-Deutsch-Cook, *Research Methods in Social Relations* (Revised). New York: Henry Holt and Company, Inc., 1959.
9. Worthen, John, *Processes in Education: Delineation, Potential and Realization*. Cambridge: Harvard University, 1963. (Mimeographed).



*The Editorial Committee for the*

ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*will be pleased to consider research articles from contributors outside the Province of Alberta, for publication in the Journal.*

*Manuscripts should be addressed to the Editor.*

H. E. Smith

# The Alberta Journal of Educational Research

Vol. X, No. 3

September, 1964



THE COMMITTEE ON EDUCATIONAL RESEARCH

*Faculty of Education*

*University of Alberta*





# VALIDITY OF REGRESSION EQUATIONS AFTER SIX YEARS TO PREDICT FRESHMAN SUCCESS IN ENGINEERING

By

D. B. BLACK

In earlier articles by the author (1960, 1961), regression equations were developed to predict freshman success in Engineering at the University of Alberta, and these equations were validated on a class of Engineering freshman a year later. The sample used to develop the equations was the 1956 entering class while the validating class was the 1957 entering class. This study seeks to test the effectiveness of these same equations after six years for the 1962 entering class of freshman Engineers, to develop new equations to meet the changing freshman curriculum, and because of the increasing problem of early university admission and the trend towards school accreditation, to examine the effectiveness of the Principal's Ratings to predict freshman Engineering success.

The sample used in this study was 102 freshman engineers who had graduated from Alberta high schools in the spring of 1962, who entered the Faculty of Engineering that fall, and for whom a full record of the seven prescribed Grade XII Departmental examination scores (English 30, Social Studies 30, Mathematics 30 and 31, Chemistry 30, Physics 30 and a foreign language 30), were available. Similarly, only those entering freshman for whom the Department of Education records indicated Principal's Confidential Ratings for the equivalent subjects used for admission to the Faculty were used in the study.

Several changes have taken place since the 1956 study. First, the Departmental scores in Mathematics 31 are now based on an examination rather than the teachers' mark or Principal's Rating. In 1957, the Faculty of Engineering replaced the 1956 Drawing 2 and 4 courses with a new course Drawing 2. Courses in Chemistry, Physics and Mathematics have been retained. Drawing 2 became Drafting and Graphics 151 while Engineering Mechanics 1 evolved into two courses, Statics (E.M. 150) and Particle Dynamics (E.M. 151). The two survey courses, C.E. 5 and 6 have been reduced to a single laboratory (field) course, C.E. 153, which was not included in the current study. The half course, Eng. P. 159, Orientation to Engineering Profession, has not been included in the current study as well.

TABLE I

COMPARISON OF 1957-58 AND 1962-63 GRADE XII PREDICTOR VARIABLE SAMPLE MEANS TO 1956-57 SAMPLE MEANS AND 1962-63 PRINCIPAL'S RATINGS TO ACHIEVED GRADES

Grade XII Predictor	1956-57 Sample N=131	1957-58 Sample N=245	1962-63 Sample (N=102)		Significance of Differences of Means		
			Ach'd Grades	Pr's Rating	1956 and 57 Samples	1956 and 62 Samples	
						Ach'd Grades	Pr's Rating
	$\bar{X}$ SD	$\bar{X}$ SD	$\bar{X}$ SD	$\bar{X}$ SD			1962 Principal's Rating and Grades Achieved
English 30 .....	65.2 10.2	63.8 9.6	67.0 9.5	64.8 10.8	NSD	NSD	NSD
Social Studies 30 .....	73.3 10.5	69.9 10.9	70.9 9.8	71.1 10.4	1%	NSD	NSD
Mathematics 30 .....	74.4 10.3	72.6 9.2	74.0 10.6	73.9 9.5	NSD	NSD	NSD
Mathematics 31 .....	69.4 14.1	70.3 11.3	68.6 11.7	68.5 10.6	NSD	NSD	NSD
Physics 30 .....	72.8 10.6	70.6 9.4	73.8 10.1	69.3 10.5	NSD	NSD	5%
Chemistry 30 .....	73.2 11.9	72.1 9.4	74.3 9.8	70.8 10.5	NSD	NSD	5%
Foreign Language 30 .....	66.2 11.6	65.7 10.1	67.2 10.7	63.3 12.3	NSD	NSD	5%
Mathematics Average .....	71.6 10.8	71.3 8.7	71.6 10.1	71.5 9.1	NSD	NSD	NSD
Science Average .....	72.7 10.4	71.1 8.5	74.0 8.9	69.9 9.1	NSD	NSD	5%

The method of the study was to use the Grade XII Departmental scores, both those actually achieved on the Departmental Examinations (Act.), and the Principal's Ratings (P.R.), in the 1956 regression equations to predict the freshman success of the 1962 sample of Engineering freshman. Where the Engineering courses or criterion variables were approximately the same as in 1956, these predicted and achieved grades in these courses were correlated. This was done on the assumption that the grading practices of the Department of Education and the Faculty of Engineering, respectively, have not changed since 1956. This assumption will be tested. The second assumption is that notwithstanding changes in the content both in Grade XII and freshman courses in six years, the skills and abilities necessary to complete a course in Mathematics, for example, will have remained unchanged, and that the scores reported from examinations are a good representation of this ability.

The second part of the study will deal with the development of new regression equations to meet the current curriculum in freshman Engineering. The Dominion Bureau of Statistics Stepwise Regression Computer Program for an IBM 1620 computer was used to develop these equations. All calculations were done at the University of Alberta's Computer Center.

### *Findings:*

To test the assumption that the grading practices of the Department and the University had not changed, the means, standard deviations were calculated for the 1962 sample for all predictor and criterion variables including the Principal's Ratings. Statistics for the 1956 and 1957 samples are those as reported earlier by the author (Black, 1961). These data and analyses are reported in Tables I and II.

Table I shows that with one exception there are no statistically significant differences between the Departmental achieved grades of the 1956 and 1962 freshman samples. Significant differences at the 5% level were found between 1962 achieved and Principal's Ratings Departmental scores in Physics 30, Chemistry 30 and Foreign Language 30, and because the significant differences in both Physics and Chemistry grades, in Science average as well. It is interesting to note that in every instance, but Social Studies 30, the mean reported Principal's Rating for the 1962 sample was lower than the actually achieved mean. In the case of Social Studies, the difference was two tenths of a mark. This consistent trend to report lower grades suggests strongly (at 1% level of confidence using the Sign Test) that average high school teacher's ratings for Engineering freshman, at least, were consistently lower than those



actually achieved. This fact plus the statistically significant differences for primary science predictors, and Physics 30 in particular, should be of great importance to the Faculty of Engineering in the event that it considers early admission based on Principal's Ratings. In terms of the conventional t-Test, there are no significant differences in overall mean high school performance of the 1956 and 1962 freshman classes. It is concluded then, that the assumption of a consistent grade admission pattern in Grade XII is a valid one.

The assumption of a consistent grading pattern in University freshman engineering grades would not appear to be as valid. These data are reported in Table II. In 1957, there appeared to be a change downward in Physics, Chemistry and Drafting marks, while mean Mathematics marks climbed. The 1962 marks tended to follow the 1957 pattern with the exception of Mathematics which dropped below the 1956 mean and which was consistent with decline in 1957 mean marks in Physics and Chemistry. Only mean marks in Engineering Mechanics remained consistent. The introduction of the English course as indicated in the 1962 sample with a mean grade of 51.6 is reflected in the consistent downward trend in Engineering Average. The Average was not significantly different between the 1956 and 1957 classes, but it was highly significant for both the 1956 and 1957 and the 1962 mean freshman performance. It can be concluded that inasmuch as there were no significant differences in the capacity of the 1962 freshman class and those entering in 1956 and 1957 as indicated by the Grade XII marks, the marking practices of the University has changed and this change would appear to have been effected primarily in the Arts and Sciences courses beginning in 1957.

Although there has evidently been a change in the level of criterion grades, this, in turn, could be viewed as a consistent change. Therefore, in correlating the predicted grades for 1957 and 1962 freshman classes, the effectiveness of the equations to predict freshman success may not be invalidated because correlation coefficients are dependent on relative ranking in a distribution rather than absolute value of the scores. The validity coefficients of the 1956 equations to predict 1957 and 1962 freshman grades are reported in Table III. Two sets of equations are used, seven variable equations using each of the seven Grade XII Departmental prerequisites for admission, and a five variable equation which uses Mathematics Average ( $\bar{X}$ ) instead of Mathematics 30 and 31 scores and Science Average ( $\bar{X}$ ) instead of Chemistry 30 and Physics 30 scores. The 1962 freshman grades are predicted using both actual Grade XII marks and the Principal's Ratings.

TABLE II  
COMPARISON OF 1956 FRESHMAN SAMPLES WITH 1957 AND 1962 FRESHMAN SAMPLES: FRESH-  
MAN MEANS GRADES

Engineering Freshman Variable	Course Number			Freshman Sample						Significance of Difference of Mean		
				1956 N=131		1957 N=245		1962 N=102				
				$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD			
	'56	'57	'62							1956 & '57 Sample	1956 & '62 Sample	1957 & '62 Sample
Engineering Average	...	...	...	63.3	10.8	61.4	11.3	57.5	13.1	NSD	1%	1%
Mathematics	2	2	202	60.5	17.7	68.0	15.9	53.1	18.5	1%	1%	1%
Physics	21	21	230	61.6	20.2	54.6	16.5	57.6	18.8	1%	NSD	NSD
Chemistry	40	40	232	60.4	14.1	57.1	13.5	55.5	12.6	5%	1%	NSD
Drafting and Graphics	2* 4*	2 2	151 151	67.2 69.8	8.8 11.2	62.1 62.1	10.8 10.8	61.3 61.3	9.7 9.7	1% 1%	1% 1%	NSD NSD
Engineering Mechanics	1 1	1 1	150 151	62.2 62.2	15.9 15.9	62.5 62.5	15.1 15.1	59.7 61.6	18.4 23.8	NSD NSD	NSD NSD	NSD NSD
Civil Engineering	5 6	5 6	...	62.5 64.6	13.0 8.2	60.9 65.8	13.5 6.8	...	...	NSD NSD	...	...
English	...	...	250	...	...	...	...	51.6	8.5	...	...	...

\*NOTE: In 1957 Drawing 4 was dropped as a course. Comparisons are made to the single Drawing Course in 1957, 1962.

TABLE III:  
ACTUAL AND ESTIMATED VALIDITY COEFFICIENTS AND  
STANDARD ERRORS OF ESTIMATE FOR PREDICTED AND  
ACHIEVED FRESHMAN ENGINEERING COURSES OF ENTERING  
CLASSES IN 1956, 1957 AND 1962 BASED ON GRADE XII  
DEPARTMENTAL GRADES AND PRINCIPAL'S RATINGS

Year	Criterion Calendar Course Number	1956 Sample (N=131)			Obtained Validity Coeffi- cients for 1957 Sample (N=245)			Obtained Validity Coefficients for 1962 Sample (N=102) Predicted from 1956 Prediction Equations			Validity Coefficients* for New Equations based on 1962 Class for 1962 Freshman Courses		
		Actual Marks			Actual Marks			Actual Marks			Actual Marks		
		7 Variable	5 Variable	R. SE.est.	7 Variable	5 Variable	R. SE.est.	7 Variable	5 Variable	R. SE.est.	7 Variable	5 Variable	R. SE.est.
	1956												
	1957												
	1962												
Engineering Average	-	.69	.67	8.0	.56	8.6		.73	6.8		.76	8.8	
Mathematics	2	.69	.68	13.0	.57	15.1		.66	10.5		.69	14.0	
Physics	21	.65	.63	15.7	.61	14.5		.63	11.4		.70	13.9	
Chemistry	40	.70	.69	10.2	.59	11.6		.73	6.5		.76	8.5	
Drafting and Graphics	2	.51	.48	7.7	.41	9.8		.61	6.0		.67	7.5	
	4	.41	.42	10.2	.42	12.4		.59	6.2				
Engineering Mechanics	1	.54	.52	13.6	.51	12.7		.49	12.0				
	1	.54	.52	13.6	.51	12.7		.66	12.8		.61	15.0	
Civil Engineering	5	.57	.55	10.9	.45	12.2		-	-		.71	17.3	
	6	.12	.08	8.2	-.12	7.6		-	-		-	-	
English	-	-	-	-	-	-		-	-		.60	7.1	
	250							-	-		.59	7.0	
											.67	10.1	
											.63	14.9	
											.58	15.8	
											.64	10.1	
											.54	8.5	
											-	-	
											.58	15.4	
											.70	17.6	
											.55	15.9	
											.62	13.4	
											-	-	
											-	-	
											-	-	
											.54	7.2	
											.55	7.5	

\*Coefficients are based on maximum R achieved by using all Grade XII predictor variables in regression equation.



The validity coefficients for the 1956 and 1957 samples have been reported earlier (Black, 1961) and will be used to compare the validity coefficients obtained for the 1962 sample. It will be noted that in the 1956 validation study using actual marks for the seven variable predictor equations, all but one of the obtained validity coefficients were lower than those for the 1957 sample. In contrast, for the 1962 sample, five of the eight validity coefficients were numerically higher when actual marks were used to predict success. The predictions using the 1962 Principal's Ratings showed that only three of eight validity coefficients were numerically higher. In every instance but two for both actual and Principal's Rating predictors, the validity coefficients obtained for the 1962 sample were numerically higher than those obtained for the 1957 sample.

The same general pattern held true for the five variable predictions. Of the validity coefficients for the 1962 sample six of eight are numerically greater than the 1956 sample; for the 1957 sample, seven of eight are equal to or higher numerically. Of the Principal's Ratings for the five variable predictions three of eight are equal to or higher numerically than the actual mark validity coefficients for the 1956 sample, and six of eight are higher than the 1957 sample.

In no instance for any comparison were differences found that were statistically significant at even the 5% level. However, the general pattern of differences favoring the 1962 sample would confirm the general observation that the regression equations developed from the 1956 sample worked equally well for the 1962 sample. The evidence does suggest a stability that was much greater than suspected and would thus favor operational implementation of the regression equations within the admission procedures of the University.

By the same token, the Principal's Ratings, as predictor variables, performed as would have been anticipated from earlier studies and further support the case for use of these marks for early provisional admission procedures.

The final part of Table III reports the validity coefficients and standard errors of measurement for new regression equations for the current freshman offering of the Faculty of Engineering. In every instance, when such comparisons can be made, the actual mark of seven and five predictor validity coefficients are equal to or numerically higher than those obtained for the 1956 sample. The validity coefficients for the equations using Principal's Ratings are slightly lower.

TABLE IV  
RELATIVE BETA WEIGHTS\* AND MULTIPLE CORRELATION COEFFICIENTS FOR SEVEN GRADE  
XII DEPARTMENTAL SCORES TO PREDICT SEVEN FRESHMAN ENGINEERING SCORES

Engineering Criterion	Grade XII Departmental Score Predictor							R	Standard Error of Estimate
	English 30	Soc. St. 30	Math. 30	Math. 31	Chem. 30	Physics 30	For. Lang. 30		
Engineering Average	-31	39	15	9	33	75	4	.762	8.8
Mathematics 202	-54	59	59	24	43	113	17	.687	14.0
Physics 230	-73	62	-11	26	101	142	14	.701	13.9
Chemistry 232	-10	32	14	4	66	47	-14	.760	8.5
Drafting and Graphics 151	-24	32	-3	7	-4	55	4	.666	7.5
Engineering Mechanics 150	-93	67	74	-14	14	138	-18	.614	15.0
Engineering Mechanics 151	-124	64	33	41	69	273	42	.713	17.3
English 250	23	25	-2	3	-1	8	2	.597	7.1

\* Multiplied by 100 to remove decimals.

As evidenced earlier by changes in the freshman Engineering curriculum and in the marking patterns of the University courses since 1956, there is need for new regression equations to replace those developed earlier. This has been done using the 1962 freshman sample of 102 cases. The computer program used in this study was such that each regression equation would be examined as each successive predictor variable in order of contribution to the prediction was added to the equation. If the program ran to its full length, i.e., all variables were used, a regression equation with a maximum multiple correlation coefficient was obtained with the criterion in which each of the variables was weighted to give an optimal relationship. It is evident from Table III, that the actual grades rather than Principal's Ratings on all seven required Grade XII Departmental subjects, afford the best predictors of freshman success at the University. Using all seven Grade XII Departmental scores, the relative weightings (Beta weights) of the different subjects for the eight 1962 freshman Engineering criteria were calculated and are reported in Table IV.

It will be noted that Physics 30 is the highest and most frequent positive contributor to the prediction of freshman success followed closely by Social Studies 30 and Chemistry 30. The two Mathematics grades do not contribute as significantly as might be anticipated. The foreign language 30 grade makes only a small contribution. Contrasting is the consistent inverse relationship between English 30 grades and all freshman grades, except English 250. Care should be taken in interpreting these data. They show only the relationships of the various predictors to the prediction of success. They do not relate in any way to the necessity of various courses as prerequisites for study. Therefore, while it is readily apparent that there is an inverse relationship between success in high school English and Engineering success, this is no way to deny the importance of English even to Engineers. The data do suggest, however, that whatever skills are being measured and presumably have been taught in Grade XII English as measured by the Departmental examinations are not those required by prospective and successful Engineering freshmen. Conversely, one is puzzled by the consistent positive contribution of Social Studies 30. It is possible that the communicative skills demanded of Engineering students are best taught and measured in this course. When one considers that even in English 250, the Social Studies 30 mark contributes about equally as the English 30 mark to the prediction of this course, the above hypothesis would seem to have some merit. What the data in Table IV do suggest is that in the selection of Engineering freshmen, the Faculty of Engineering could do well to



TABLE V  
OPERATIONAL REGRESSION EQUATIONS TO PREDICT FRESHMAN ENGINEERING GRADES  
USING SEVEN GRADE XII DEPARTMENTAL EXAMINATION SCORES AS PREDICTOR VARIABLES

Engineering Criterion	Type of Score	Regression Equations							Standard Error of Estimate	R	
		Grade XII Predictors									
		English 30	Soc. St. 30	Math. 30	Math. 31	Chem. 30	Physics 30	For. Lang. 30			
											Constant
Engineering Average	Act. P.R.	—23	.32 .19	.26	.28	.29	.70 .37	...	—22.4 —19.0	8.8 10.1	.75 .66
Mathematics 202	Act. P.R.	...	.30	.45 .58	.41	...	.74 .33	...	—55.7 —40.1	14.1 14.9	.67 .61
Physics	Act. P.R.	—38	.35	...	.40	.55 .34	.85 .50	...	—45.4 —28.5	13.9 15.7	.70 .57
Chemistry 232	Act. P.R.	...	.21	...	.31	.50 .53	.45	...	—30.0 —31.9	8.5 10.2	.75 .60
Drafting and Graphic 151	Act. P.R.	—21	.32 .19	...	...	...	.52 .39	...	14.0 20.5	7.5 8.4	.66 .53
Engineering Mechanics 150 (Statics)	Act. P.R.	—47	.35	.37	.41	...	.75 .60	...	—16.2 —10.1	15.0 16.0	.61 .51
Engineering Mechanics 151 (Particle Dynamics)	Act. P.R.	—35	...	.47	.34 .60	.41	1.19 .59	...	—56.3 —55.0	17.3 19.3	.70 .60
English 250	Act. P.R.	.25 .19	.29 .20	...	...	...	.11 .13	...	5.3 16.2	7.1 7.1	.59 .53

ask only for passing grades in both English 30 and a foreign language 30 and exclude these marks from any inclusion in the calculations that determine whether a freshman will or will not be admitted to the Faculty.

It is also apparent from earlier studies that prediction equations based on all predictor variables can be successfully reduced with little loss in predictive efficiency. This has the advantage of making the predictions much easier to calculate in the high schools which do not normally have access to electric calculators or computer services. To reduce the number of predictor variables, the relative beta weight of each predictor added successively to the equation was tested for its significance greater than zero. When this hypothesis (Null hypothesis) could be accepted at the 10% level of significance, subsequent predictor variables by virtue of the programs selection of variables in order of their contribution to the prediction were dropped. These "reduced" regression equations are identified as "operational" regression equations and are reported in Table V ready for use in the schools. Equations have been calculated for both actual marks and Principal's Ratings. The five predictor variable equations are not reported because of their limited use due to the prescriptive nature of the prerequisite Grade XII scores for admission. The seriousness of the reduction from seven to three or four predictor variables is readily apparent by comparing the multiple correlation coefficient and standard errors of estimate in Table IV to those for actual grades in Table V. In no case was the multiple correlation coefficient lowered by more than two points or the standard error of estimate increased by more than two tenths of a score point.

### Summary

This study has sought to evaluate the effectiveness of regression equations developed in 1956 to predict success for a freshman engineering class six years later. It was found that they remained effective over this period. Moreover, when Principal's Ratings were used instead of marks actually obtained on the Grade XII Departmental examinations, the predictions remained sufficiently high to be useful for admission procedures. The findings offer promise for their operational incorporation in the admission practices of the Faculty and have implications for inaugurating a practice of early (Easter) provisional admission to the Faculty pending successful completion of the June Departmental examinations. Certainly the stability of the equations which was found to be higher than would have been suspected, offers further encouragement in this direction.

New regression equations have been developed to accommodate changes in the curriculum and grading practices of the freshman year. It has been shown that those equations can be modified with little loss in predictive efficiency to make them more practical for use by school counsellors.

#### BIBLIOGRAPHY

Black, D. B. "The Prediction of Freshman Success in the University of Alberta from Grade XII Departmental Results." *Alta. Journal of Ed. Res.*, VI:1 (March, 1960), 30-53.

"Differential Grade Prediction: A Note on Its Effectiveness for a Freshman Class of Engineers." *Alta. Journal of Ed. Res.*, VII:2 (June, 1961), 86-92.



AN INVESTIGATION INTO THE KNOWLEDGE OF  
MATHEMATICS OF ALBERTA STUDENTS, BASED ON THE  
GRADE NINE DEPARTMENTAL EXAMINATION WRITTEN  
IN JUNE, 1960

ELLEN A. LAWS

SCORES: GRADE IX EXAMINATIONS

Each year in June the grade IX high school entrance examinations are conducted under the direction of the High School Entrance Examinations Board. Students write papers in language, literature, social studies, science, mathematics, and reading comprehension. They pass or fail on the basis of their achievement on these papers. In addition they write a test of general aptitude.

One of the major problems of the Examinations Board is that of setting standards for judging satisfactory achievement. The present practice of scaling final marks has received much criticism on the basis that standards have fallen in recent years. The writer is not prepared to judge the accuracy of this contention, but it must be remembered that economic pressure is forcing students to remain in high school who, twenty or thirty years ago would have dropped out at the end of grade VIII and found gainful employment. The result is that students of all but the lowest levels of academic ability are writing grade IX examinations.

The Royal Commission on Education investigated the field of departmental examinations (1). It found that in the years from 1906 to 1935, the marks awarded to a student depended primarily upon the difficulty of the examination. Passing marks based on raw scores, varied from year to year. Various systems of bonus marks were used to achieve a reasonable distribution of marks, but lack of consistency in their application rendered the system unsatisfactory.

The present policy of the Department of Education is as follows:

The grading of the pupils is based on the assumption that in a LARGE group the students from year to year are about the same in ability and achievement, that is to say, from one year to the next there is the same percentage of average, of dull and of bright pupils. It would follow that the same proportion of pupils should receive high marks each year and that the percentage of failures should not vary. The actual proportion of pupils in each category was set several years ago, and except for occasional minor changes, is used from year to year and applies to almost all of the subjects (2).

The following table shows the distribution of letter gradings which was used for grade IX pupils in 1960 when this study was begun\*—

TABLE I  
DISTRIBUTION OF LETTER GRADINGS  
TO GRADE IX PUPILS

Letter gradings given to pupils	Percentage marks used to derive letter gradings	Percentage of total number of pupils
H .....	80-100	10
A .....	65- 79	25
B .....	50- 64	35
C .....	40- 49	20
D .....	0- 39	10

\*The present distribution became effective in 1962.

When departmental examinations are marked a raw score for each paper is established. This is later converted to a scaled score. The student with the highest raw score for the province is given a scaled score which is usually one hundred percent, while the student with the lowest raw score is assigned a scaled score near zero. The raw scores are arranged in rank order. The top ten percent are equated to percentage gradings ranging from one hundred percent down to eighty percent by means of a linear graph. These students are given an H grading. The same method is used in equating percentage scores and gradings for other categories.

While percentage scores are kept on record at the Department of Education, only letter gradings appear on the diplomas sent to the individual pupils. To the teachers receiving these pupils in grade X, the letter gradings and the percentage marks which they represent indicate only the relative position of the pupil to the entire student body of the province. A B pass mark, for instance, which is shown on the statement issued to the pupils as indicating a score of from fifty to sixty-four percent, does not mean that the raw score was necessarily in the fifty to sixty-four range, or that the pupil knew at least one-half of the material tested on the examination paper. It seems desirable for grade X teachers to have a fairly clear picture of their pupils' ability in relation to the grade IX course, as well as their ability in relation to the other pupils in the province. It was with this object in mind that the writer launched an investigation into the grade IX Mathematics papers written in the year 1960.

SCORES: GRADE IX MATHEMATICS EXAMINATIONS

The data for the study came from a proportional systematic sample of 389 papers, selected with the help of the Examinations Branch from the total of all Grade IX mathematics papers written in the Province of Alberta in June, 1960. The sample contained the same proportion of papers in each of the categories, H, A, B, C, and D as the diplomas issued in the province that year.

The first step in the investigation of the sample papers was the preparation of an item analysis. The principles involved in the solution of each examination question were itemized and these were used as the basis for scoring the sample papers.

The grade IX course in mathematics includes four areas of study, arithmetic, algebra, geometry, and graphic study which is the visual presentation and interpretation of numerical data. A separate frequency distribution was made for each category of student, H, A, B, C, D, in each of the four areas of study. Each item was checked according to whether the principle was correct, incorrect or not attempted. In solutions involving more than one principle, each principle was examined independently of those that preceded or followed it, and credit was given for knowledge of that principle without regard to the values which had been assigned when the paper was received. Where the same principle occurred more than once on the examination paper, the totals were combined.

The performance of each category of student, H, A, B, C, D, was computed separately for each of the four areas of study. The results are summarized in Tables II-VI. To arrive at the figures shown in the tables, for each category of student the total number of correct responses in a single area of study, e.g. arithmetic, was expressed as a percentage of the total possible in the area. For comparison, each table also includes the percentage implied by the letter grading issued to the student.

TABLE II  
RELATIVE COMPETENCE OF THE H GROUP OF  
STUDENTS IN THE FOUR AREAS OF STUDY

H (80 - 100 percent)		
Arithmetic .....	78.51	per cent
Algebra .....	86.74	per cent
Geometry .....	81.36	per cent
Graphic Study .....	83.97	per cent



TABLE III

## RELATIVE COMPETENCE OF THE A GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

A (65 - 79 percent)		
Arithmetic .....	60.19	per cent
Algebra .....	67.75	per cent
Graphic Studies .....	60.29	per cent
Geometry .....	77.12	per cent

TABLE IV

## RELATIVE COMPETENCE OF THE B GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

B (50 - 64 percent)		
Arithmetic .....	43.09	per cent
Algebra .....	46.33	per cent
Geometry .....	44.23	per cent
Graphic Study .....	60.83	per cent

TABLE V

## RELATIVE COMPETENCE OF THE C GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

C (40 - 49 percent)		
Arithmetic .....	24.99	per cent
Algebra .....	23.77	per cent
Geometry .....	30.91	per cent
Graphic Study .....	42.95	per cent

TABLE VI  
RELATIVE COMPETENCE OF THE D GROUP OF  
STUDENTS IN THE FOUR AREAS OF STUDY

D (0 - 39 percent)		
Arithmetic .....	15.04	per cent
Algebra .....	10.53	per cent
Geometry .....	17.96	per cent
Graphic Study .....	29.73	per cent

### CONCLUSIONS

It must be made clear that any conclusions drawn in this study are founded entirely on the results of one Departmental examination. Therefore, they must be qualified with the observation that this is what the students knew, or did not know, in 1960. In no way is this investigation intended to establish final conclusions regarding the ability of all students who have written, or will continue to write the grade IX examination.

However, the unsatisfactory standard of work exhibited by the B, C, and D groups in this investigation poses the same question that has been asked by other investigators—"What can be done to raise the level of achievement in mathematics?" Is the answer to be found only in improved instructional procedures, in more intensive drill on mechanical processes, and in stricter discipline in mathematical routines? Or should we take a second look at the mathematics curriculum and ask ourselves whether our present course is designed to do what it is intended to do. Mathematics, to justify its place in the school curriculum, must contribute to the satisfaction of the needs of the student in the fields of personal living, and social and economic relationships.

### REFERENCES

1. Report on the Royal Commission on Education. Edmonton: Government of Alberta, The Queen's Printer, 1959.
2. Department of Education, Province of Alberta. "Curriculum News Letter, No. 9." Edmonton: The Queen's Printer, February, 1958.

# REFERENCE GROUP BEHAVIOR AMONG HIGH SCHOOL STUDENTS

HENRY ZENTER

*University of Alberta, Calgary*

This paper examines certain aspects of reference group behavior among high school students of varied racial, socio-economic and ecological description in both Oregon and Alberta.

Although the concept of reference group behavior was recently formulated by Merton, the idea that an individual is influenced in his behavior by other persons and groups that are in some sense meaningful to him has long been held as a fundamental principle of behavior among sociologists and social psychologists.<sup>1</sup> George H. Mead, in particular, attributed central importance to 'significant others' in the socialization of the child.<sup>2</sup> Thus, while the underlying notion of influence upon behavior emanating from others has led to extensive speculation and considerable research in sociology in recent years, it would, no doubt, readily be agreed that we are far from an adequate understanding of the circumstances under which given individuals will take given categories of 'significant others' into account. Neither is it known to what extent 'significant others' influence the behavior of given individuals under given circumstances. Perhaps least of all among high school students.

The present inquiry, therefore, is focussed upon the relationship between the high school students and a number of selected 'significant other' reference categories respecting two major role-taking decisions facing students.<sup>3</sup> The two role-taking situations selected for study include: (1) graduating from high school; and (2) going on to further training after high school has been completed. The data reported here are taken from a series of more

1. Robert K. Merton, *Essays in Social Theory and Social Structure*, rev. ed., New York. The Free Press of Glencoe, Inc., p. 299 ff.
2. George H. Mead, *Mind, Self and Society*, University of Chicago Press, 1934.
3. Ralph H. Turner, "Role-Taking, Role Standpoint, and Reference-Group Behavior", *American Journal of Sociology*, LXI (1956), 316-328. Turner identifies three major types of reference groups that lie on a continuum. Following Sherif and Shibutani, he places the *identification group* at one end of the continuum since it is seen as the source of values with the result that the individual adopts the member's standpoint as his own. At the opposite end of the continuum is a second type called the *interaction group*, so designated since the members of such a group constitute merely conditions to the behavior of the actor. In between is a third type of group called the *valuation group*. The latter is differentiated from the remaining two types in that its effect upon the individual's behavior will be determined by the valuation which his more basic orientation leads him to place upon it. It is the latter type which would appear to subsume the reference categories employed in the present study. Included are: relatives, teachers, school friends, friends of the family, other people liked, and neighbors. Since there is no theoretical justification for the assumption that these classes of persons constitute real groups, the term "reference category" has been employed in the present discussion.



comprehensive attitudinal surveys among high school students in Oregon and Alberta made during the past three years.

### THE SEVERAL POPULATIONS

Involved in the series of surveys were both Indian and non-Indian students attending an integrated small-town high school in Central Oregon, the identity of which must for the present remain anonymous. Alberta students also involved include: (1) Blood Indian and non-Indian students attending the partially integrated small-town high school at Cardston, Alberta; (2) Blood Indian students attending St. Mary's residential school on the Blood Reserve near Cardston; (3) Blackfoot Indian students attending a partially integrated small-town high school at Gleichen, Alberta; (4) Blackfoot Indian students attending the Chief Crowfoot residential school on the Blackfoot Reserve near Cluny, Alberta; and (5) non-Indian students attending high school in an upper middle class residential area in the city of Calgary, Alberta. For convenience, these several populations have been grouped into five categories: (1) Oregon Indians; (2) Oregon Whites; (3) Alberta Indians; (4) Rural Alberta Whites; and (5) Urban Alberta Whites.

The Oregon data were obtained from populations consisting of 52 Indian and 304 non-Indian students. The Alberta data were obtained from populations consisting of 115 Blood and Blackfoot Indians, 335 rural non-Indian students, and 193 urban non-Indian students. For the most part, these totals represent the usable returns obtained from a questionnaire which was circulated in the classroom among the high school populations in each of the several schools involved.<sup>4</sup> In those instances where the entire student body was surveyed, the net response rate was in no case below 80 per cent.

In terms of basic composition the five populations showed considerable difference in respect to sex, age, class standing, rural-urban residence and father's education.

Differences in sex composition were minimal, with males predominating slightly over females in all instances except the Oregon Indian population in which the proportion of the sexes were approximately six to four in favor of the females. With respect to age, both the Oregon and the Alberta Indian populations showed considerably higher average ages than the three non-Indian populations. In terms of class standing, both Indian populations tended to bulk larger in grades IX and X, with markedly smaller numbers in

4. In two of the five populations, some variation in data collection procedure obtained. Nine of the 115 Alberta Indians were polled at the denominational residence in which they were domiciled while attending a partially integrated public high school in the town of Gleichen. Alternatively, the 193 Urban Alberta non-Indian students were selected primarily on the basis of convenience and availability during a library period on a given day. Consequently neither the degree of numerical nor the social representativeness is fully known.

grades XI and XII, as compared with a much more uniform distribution over the four grades among the two small-town non-Indian populations. Those attending the city high school were drawn from grades X and XI exclusively.

In regard to rural-urban differences in residence, the respective proportions residing on farms were approximately equal, with the exception of the Alberta Indian population in which the relative proportion was twice as large as among the remaining three small-town high school attending populations. Further, only a small proportion of both Indian populations lived in either a town or a city, as compared with more than half of the students in both non-Indian small-town populations. Those attending the city high school were, of course, almost all urban residents.

Respecting father's educational achievement there were again marked differences. The two Indian populations had a much larger proportion of students whose fathers completed only elementary school or less than did the three non-Indian populations. Also, more Indian students gave a "don't know" response on the question pertaining to father's education than did non-Indians.

### **The Data**

It will be convenient to present the data under sub-headings which bear upon the two types of role-taking situations facing high school students. Further, under each of these two sub-headings data will be presented which correspond to the assessed reaction of others and the importance attributed to such reaction by the students themselves. With respect to both types of decisions, the 'significant other' categories which students were asked to evaluate were identical and included the following: (1) parents; (2) school friends; (3) relatives; (4) friends of the family; (5) neighbors; (6) teachers; and (7) other people liked (you like).

**Graduation from High School:** Two questions were put to the respondents concerning the first of the two role-taking decisions, namely, graduation from high school. The first question referred to the reaction of parents and was phrased as follows: "How disappointed in you would your parents or guardians be if you didn't finish high school?". Answer categories provided were four in number, including "very", "somewhat", "very little", and "not at all". In order to facilitate comparison of the responses to this question with that which follows, those students who reported that their parents would be "very" or "somewhat" disappointed were combined into a single category referred to in Table I as "High Parental Disappointment". Conversely, those students who reported that their parents would evidence "very little" or "no" disappointment are also combined and referred to in Table I as reporting "Little or No Parental Disappointment".



TABLE I  
FREQUENCY OF ASSESSED NEGATIVE REACTION ON  
THE PART OF PARENTS RESPECTING FAILURE TO  
COMPLETE HIGH SCHOOL TRAINING

Response Category	Populations				
	Alberta Indians	Rural Alberta Whites	Oregon Indians		
	N=115	N=335	N=52		
	Oregon Whites	Urban Alberta Whites			
	N=304	N=193			
	A.I.	R.A.W.	O.I.	O.W.	U.A.W.
	%	%	%	%	%
High Parental Disappointment .....	90.4	97.0	98.0	99.0	96.3
Little or no Parental Disappointment .....	9.5	2.9	2.0	1.0	3.6
A.I.	=	Alberta Indians			
R.A.W.	=	Rural Alberta Whites			
O.I.	=	Oregon Indians			
O.W.	=	Oregon Whites			
U.A.W.	=	Urban Alberta Whites			

The data reported in Table I show that for each of the five populations the proportions of students reporting high parental disappointment exceeds 90 per cent. Moreover, the range of variation between populations is quite restricted, suggesting that the assessed reaction of parents to the student's potential failure to complete high school is independent of differences in nationality, socio-economic and ecological status.

The second question which was presented to the respondents referred to the reaction of the remaining 'significant other' categories respecting failure to complete high school. The wording of the question was as follows: "Aside from your parents or guardians, would any of these people be disappointed in you if you didn't finish high school?". The response categories provided permitted only a simple "yes" or "no" answer, although in each case a small proportion of the respondents gave no answer to the question. The data reflecting the relative strength of the assessed reaction as measured by frequency of affirmative response of these remaining categories of 'significant others' are shown in Table II.

Table II shows that the students expected far less frequently than was the case with their parents that the remaining categories of 'significant others' would be disappointed at their failure to complete high school. A comparison of Table II with Table I shows the





proportions reporting that their relatives (the category with the highest frequency of reported disappointment) would be disappointed are in each instance smaller than that reporting parental disappointment. It is noteworthy, however, that among these 'significant other' categories relatives uniformly occupy the position of greatest importance, while neighbors occupy the position of least importance. Even more striking, perhaps, is the median to low position of the category of school friends. Table II shows, further, that the rank order of the categories is identical for Alberta Indians and Rural Alberta Whites, with relatives occupying first position, followed by teachers, school friends, friends of the family, other people liked, and finally, neighbors. In the case of the Oregon Indians the response pattern is roughly the same with the exception of the teacher category which drops from second to fifth place in rank. Among Oregon Whites the response pattern is again similar with the exception of the category "other people liked" which occupies third instead of fifth position in the rank order. Among Urban Alberta Whites the ranking shows considerable variation with "other people you like" ranking in second place, followed by friends of the family in third position, and teachers and school friends being of equal significance and occupying fourth and-or fifth place, respectively.

Turning now to the issue of the importance attributed by students to the assessed reaction of 'significant others', it is unfortunately the case that data are available only for the Urban Alberta White population. The question put to these students was phrased: "Will your decision to finish or not finish high school be influenced to any extent by the fact that the following people might be disappointed in you if you didn't finish?". The categories presented for evaluation omitted parents, but did include school friends, relatives, friends of the family, neighbors, teachers, and other people liked. Again, the response categories called for a "yes" or "no" answer and a small proportion of the students made no answer to the question. These data are presented in Table III.

Reading in the percentage affirmative column of Table III it is apparent that once again relatives are most frequently cited as influential in making a decision concerning completion of high school and neighbors least frequently. The category "other people you like" ranks in second place followed by teachers in third position. School friends and friends of the family occupy fourth and fifth position, respectively. Once again it is perhaps noteworthy that the category school friends occupies a median to low position in the rank order.

TABLE III

CATEGORIES OF REFERENCE RANKED IN THE ORDER  
OF FREQUENCY OF INFLUENCE RESPECTING THE  
DECISION TO COMPLETE HIGH SCHOOL TRAINING

Reference Category	Population		
	Urban Whites N=193		
	Percentage	Percentage	Percentage
	Affirmative	Negative	No Answer
	U.A.W.	U.A.W.	U.A.W.
	%	%	%
Relatives .....	59.0	36.2	4.6
Other People Liked .....	39.9	51.3	8.8
Teachers .....	37.8	51.3	10.8
School Friends .....	32.6	59.5	7.7
Friends of the Family .....	30.5	60.6	8.8
Neighbors .....	9.8	77.7	12.4
U.A.W. = Urban Alberta Whites			

**Further Training After High School:** Two essentially parallel questions were put to the respondents concerning the assessed reaction of others with respect to the role-taking decision to go on to further training after completion of high school. The first question was worded: "If you don't decide to go on to further training (after completing high school), how disappointed in you will your parents or guardians be?". The response categories provided were the same as those used in connection with the question regarding parental disappointment at the prospect of failure to complete high school, and they were treated in Table IV in the same manner as those shown in Table I.

Table IV shows that while students very frequently reported high parental disappointment, the range of variation is somewhat greater than that observed on the parallel set of data reported in Table I. Nevertheless, it will be evident upon examination of Table V below that the frequency of assessed disappointment on the part of parents in each case substantially exceeds that reported by students for other categories of 'significant others'.

Table V shows the data based upon the question concerning assessed reaction of the same list of 'significant other' categories to the students failure to go on to further training after high school. The specific phrasing of the question was: "Aside from your parents or guardians, will any of these people be disappointed in you if you don't go on to further training (after high school)?" Once again, the percentage affirmative column shows that relatives uniformly



occupy the highest rank and neighbors the lowest. Similarly, among Alberta Indians, Rural Alberta Whites, and Oregon Whites teachers occupy second place in the rank order, while among Oregon Indians and Urban Alberta Whites teachers occupy fourth and third position, respectively. The category "school friends" again occupies a median to low position in the rank ordering as do the remaining categories.

Turning again to the issue of the importance attributed by students to the assessed reaction of 'significant others' respecting the decision to go on to further training, there is once again data available only for the Urban Alberta population. In this instance the question put to these students was worded: "Will your decision to go on to further training or not go on be influenced to any extent by the fact that the following people might be disappointed in you if you didn't go on to further training?". The categories of 'significant others' to which the students were asked to make response is identical with those listed in the previous question. The results are given in Table VI.

TABLE IV  
FREQUENCY OF ASSESSED NEGATIVE REACTION ON THE  
PART OF PARENTS RESPECTING FAILURE TO GO ON TO  
FURTHER TRAINING AFTER HIGH SCHOOL

Response Category		Populations				
		Alberta Indians	Rural Alberta Whites	Oregon Indians		
		N=115	N=335	N=52		
		Oregon Whites		Urban Alberta Whites		
		N=304		N=193		
		A.I.	R.A.W.	O.I.	O.W.	U.A.W.
		%	%	%	%	%
High Parental						
Disappointment .....		93.9	94.9	84.0	71.3	90.1
Little or no Parental						
Disappointment .....		6.1	5.1	16.0	28.7	9.9
A.I.	=	Alberta Indians				
R.A.W.	=	Rural Alberta Whites				
O.I.	=	Oregon Indians				
O.W.	=	Oregon Whites				
U.A.W.	=	Alberta Urban Whites				

TABLE V  
CATEGORIES OF REFERENCE RANKED IN THE ORDER OF FREQUENCY OF ASSESSED  
NEGATIVE REACTION RESPECTING FAILURE TO GO ON TO FURTHER TRAINING AFTER  
HIGH SCHOOL

Reference Category	Populations														
	Alberta Indians N=115			Rural Alberta Whites N=335			Alberta Urban Whites N=193			Oregon Indians N=52					
	Oregon Whites N=304			Percentage Affirmative			Percentage Negative			Percentage No Answer					
	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %
Relatives .....	75.6	71.6	67.3	63.3	60.6	16.5	22.6	28.8	33.0	32.6	7.8	5.6	3.8	3.6	6.7
Teachers .....	71.3	69.5	32.6	62.3	36.7	13.0	22.3	32.6	30.3	49.2	15.6	8.0	34.6	7.3	13.9
School Friends .....	53.9	37.9	32.6	35.0	21.7	30.4	52.8	40.3	56.6	65.2	15.6	9.2	26.9	8.3	12.9
Friends of Family .....	49.9	49.2	44.2	38.0	34.7	32.1	40.9	32.6	53.0	51.8	20.8	9.6	23.0	9.0	13.4
Other People You Like ....	41.7	42.0	42.3	44.3	43.0	34.7	46.2	32.6	46.6	43.5	23.4	11.6	25.0	9.0	13.7
Neighbors .....	14.7	27.7	11.5	23.3	15.5	53.9	61.4	59.6	65.0	67.3	31.3	10.7	28.8	11.6	17.1
A.I. =	Alberta Indians														
R.A.W. =	Rural Alberta Whites														
O.I. =	Oregon Indians														
O.W. =	Oregon Whites														
A.U.W. =	Urban Alberta Whites														

TABLE VI

CATEGORIES OF REFERENCE RANKED IN THE ORDER OF  
FREQUENCY OF INFLUENCE RESPECTING THE DECISION  
TO GO ON TO FURTHER TRAINING AFTER SCHOOL

Reference Category	Population		
	Urban Alberta Whites N=193		
	Percentage Affirmative	Percentage Negative	Percentage No Answer
	U.A.W. %	U.A.W. %	U.A.W. %
Relatives .....	56.4	36.7	6.7
Other People Liked .....	33.1	55.9	10.3
Teachers .....	31.0	57.5	11.4
Friends of the Family .....	26.9	61.6	11.4
School Friends .....	21.7	67.8	10.3
Neighbors .....	8.2	79.2	12.4
U.A.W. = Urban Alberta Whites			

Reading in the percentage affirmative column of Table VI, the now familiar pattern of response is once again apparent. Relatives uniformly occupy the position of greatest influence with respect to the decision to go on to further training, while neighbors occupy that of least importance. Again, the category "other people you like" occupies second place, and teachers third place in the rank ordering of the categories. Finally, friends of the family appear in fourth place, while school friends appear in fifth place, being slightly more influential than neighbors.

DISCUSSION

Despite modern rural-urban trends and cultural contact the finding that relatives uniformly occupy the position of greatest importance and neighbors that of lowest in each of the three role-taking situations was somewhat unexpected. Even more surprising, however, was the median to low position occupied by the category of school friends. The recent postulation by Coleman of a more or less independent adolescent culture with values antithetical to the culture of adult society finds little support in the data here reported.<sup>5</sup> Although the ecological conditions which characterize the populations studied no doubt differ somewhat from those studied by Coleman, the present findings suggest that adolescents respond selectively and situationally to the values of the social milieu in

5. James S. Coleman, "The Adolescent Subculture and Academic Achievement", *American Journal of Sociology*, LXV (1960), 337-347.



which they move.<sup>6</sup> The overall uniformity of response pattern among the five different populations in purview, however, lends additional support to this conclusion.

Clearly, the several categories of 'significant others' appear to be taken into account differentially by the students, depending upon the specific nature of the decision facing them. Certain categories of 'significant others' appear to be seen as more concerned with some kinds of role behavior than they are with respect to other kinds. Whether the fact that parents and relatives occupy the position of most extreme assessed reaction and neighbors the least extreme position is a result of differential information or differential identification or both cannot be determined directly from the nature of the data. Similarly, the medium to low position occupied by the category "school friends" raises the question as to whether they have lower visibility, in the sense that there is less information available to the respondent concerning their disposition, or whether they are adjudged to have different value standards in comparison with alternative categories of 'significant others'. The data, nevertheless, strongly suggest that the students' assessments of the reactions of the several categories of 'significant others' is predicated upon both imputed interest and variegated value standards. Moreover, both are seen as relative to the role-taking issue in purview.

Such an interpretation is warranted by the fluctuation in the relative strength of assessed reaction of the various categories of 'significant others' in the different role-taking situations. Relatives, for example, appear to be regarded as more intensely involved respecting the issue of high school graduation than they are in respect of further training. Correspondingly, the fluctuations in rank order of assessed reaction of the remaining categories of 'significant others' and the degree of importance attributed to such reactions in respect of the different role-taking situations would appear to be accountable on the basis of this same premise.

6. Such an interpretation closely coincides with conclusions reached by Elkin and Westley in their study of adolescent behavior in a Montreal suburb. See Fredrick Elkin and William Westley, "The Myth of Adolescent Culture", *American Sociological Review*, 20 (1955), 680-684.

# TOWARDS THE IMPROVEMENT OF METHODS COURSES AND PRACTICE TEACHING

MERRON CHORNY AND ARTHUR G. STOREY

*Faculty of Education  
University of Alberta, Calgary*

A methods course-practice teaching study<sup>1</sup> conducted at the university in Calgary during the 1961-62 academic year failed to disclose any significant relationship between the marks earned in an English methods course and those awarded in practice teaching. The study did, however, discover a positive and significant correlation between changes in student attitudes towards English and skill in teaching. Utilizing the cues provided by this earlier study the present authors set out to improve both the experiences and attitudes afforded a group of students through an English methods course and practice teaching.

The twenty-eight students registered at the university with a teaching major in English were selected as subjects for the experiment. The subjects were divided randomly into two equal groups and assigned to the present authors as student advisees. A committee consisting of the principals of the Elementary and Junior High demonstration schools and the head of the demonstration High School's English Department together with the present authors was organized to delimit the goals to be attempted and outline the procedures designed to achieve them.

In general the aims included<sup>2</sup>:

1. The broadening of practice teaching experiences to include the total demonstration school plant, staff, and student body.
2. The integration of student teachers and teachers with a view to making the students feel they were accepted members of the school staff and the teaching profession.
3. As close a liaison as possible between the English methods course and the experiences provided through practice teaching.
4. A resolution that staff advisors make themselves as available as possible for consultations with their advisees and that their relationships with these advisees be permissive and helpful rather than judgmental.
5. The instigation of seminars involving the members of the committee, demonstration teachers and student teachers following each of the three rounds of practice teaching. The purpose of these seminars was the gathering of opinions in order to obtain feedback for the improvement of both practice teaching and the English methods course.

One of the staff advisors for the experimental group was currently offering the English methods course; the other advisor offered it the previous year. The grades awarded in both the methods course and practice teaching were determined in the same way as for students in general.

## Instrument

A forty-five item questionnaire was devised to measure the degree to which each of the above aims was achieved. TABLE I gives the items measuring each aim.

TABLE I  
QUESTIONNAIRE ITEMS MEASURING EACH AIM

Aim	Items
1	1- 5
2	6-15
3	16-27
4	28-37
5	38-45

At the conclusion of practice teaching the questionnaire was administered to the experimental group together with a control group matched with them for age, sex, educational standing, marital status, and if married the number and ages of their children.

In part the measuring instrument reads:

Respond to each of the following statements as frankly and as honestly as you can. Do not change any response once you have made it. We want your first and most frank reaction. If a statement is Entirely True of your own practice teaching experience circle the ET category. If it is

- Mostly true circle the MT
- Somewhat true circle the ST
- Neither true nor false circle the U
- Somewhat false SF
- Mostly false MF
- Entirely false EF

Respond to each statement in turn. Do not skip any. Do not put your name on the paper but give us your honest opinion.

STATEMENT

OPINION

I was able to familiarize myself with the schol's disciplinary methods ET MT ST U SF MF EF

Procedure

By awarding the U category a weighting of zero and those out from it in either direction weights of 1, 2, and 3 plus or minus depending on the direction both positive and negative scores were obtainable. For example, the scoring for the particular item given above was:

	ET	MT	ST	U	SF	MF	EF	Total
No. of control subjects choosing each category	5	1	10	2	5	5	0	28
Score	+15	+2	+10	0	-5	-10	0	+12
	ET	MT	ST	U	SF	MF	EF	Total
No. of exp. subjects	2	10	8	5	2	1	0	28
Score	+6	+20	+8	0	-2	-2	0	+30

The Chi-square calculated on the hypothesis of equal probability, i.e. the hypothesis that there will be no difference between the total score earned by the two groups, is 7.72, which, with one degree of freedom, is significant above the .01 level. The null hypothesis must therefore be rejected and the conclusion drawn that the experimental and control groups had differential experience insofar



as familiarization with the demonstration schools' disciplinary practices was concerned.

TABLE II summarizes the comparative scores of the two groups on all items.

TABLE II  
EXPERIMENTAL AND CONTROL GROUP SCORES

Item	Exp. Gp.	Con. Gp.	D	Chi <sup>2</sup>	Sig.
1	43	40	3	.11	Nil
2	45	— 2	47	51.40	.01
3	18	— 6	20	48.00	.01
4	44	37	7	.61	Nil
5	30	12	18	7.60	.01
6	— 1	6	7	7.20	.01
7	9	17	8	2.46	Nil
8	24	37	13*	2.77	Nil
9	30	29	1	.01	Nil
10	35	35	0	.00	Nil
11	40	42	2*	.05	Nil
12	13	5	8	3.56	Nil
13	50	49	1	.01	Nil
14	32	31	1	.01	Nil
15	27	20	7	1.04	Nil
16	41	—31	72	518.40	.01
17	34	—20	54	208.30	.01
18	31	— 9	40	72.73	.01
19	37	—21	58	210.25	.01
20	53	15	38	21.19	.01
21	18	—23	41	336.20	.01
22	23	7	16	8.53	.01
23	44	18	26	10.90	.01
24	73	17	56	34.40	.01
25	60	7	53	41.80	.01
26	32	—17	49	160.07	.01
27	4	13	9*	4.76	.05**
28	45	55	10	1.00	Nil
29	63	57	6	.30	Nil
30	56	42	14	1.92	Nil
31	43	38	5	.31	Nil
32	53	45	8	.65	Nil
33	36	28	8	1.00	Nil
34	49	36	13	1.99	Nil
35	40	36	4	.21	Nil
36	56	53	3	.08	Nil
37	33	35	2*	.06	Nil
38	49	42	7	.54	Nil
39	44	22	22	7.33	.01
40	19	4	15	9.78	.01
41	24	10	14	5.77	.05
42	41	27	14	2.88	Nil
43	34	29	5	.39	Nil
44	47	47	0	.00	Nil
45	41	40	1	.01	Nil

Mean 36.67 20.86  
S.D. 14.80 22.29  
t. of Mean D 3.92 (Sig. .01 level)

\* items favoring the control group

\*\* items favoring the control group significantly

TABLE III gives the final marks in practice teaching along with those in English methods.

TABLE III  
PRACTICE TEACHING AND ENGLISH METHODS MARKS

Subject	Practice Teaching Mark	English Methods Mark
1	75	65
2	63	61
3	67	62
4	75	70
5	67	41
6	63	47
7	63	47
8	80	63
9	69	65
10	69	60
11	65	65
12	70	80
13	63	60
14	77	74
15	68	68
16	65	56
17	67	61
18	85	71
19	68	81
20	60	51
21	69	66
22	63	84
23	71	55
24	70	72
25	68	55
26	85	80
27	72	73
28	68	57
Mean	69.25	63.79
S.D.	6.59	9.48
Rho coef. of corr.	.54	

### Discussion

An examination of TABLE II indicates that aims one and three were reached within the definition of the questionnaire items. Aims two, four and five were not achieved except as sampled by items 6, 39, 40, and 41. It is to be noted that item 27 discriminated between the two groups at the .05 level but in favor of the control group. The  $t$  of 3.92 reported in TABLE II indicates that there was an overall enrichment of the experiences offered in both practice teaching and methods course where the experimental group was concerned.

The coefficient of correlation of .54 reported in TABLE III indicates that there was a positive and significant relationship between practice teaching and methods course marks where no significant relationship existed a year earlier.

### Recommendations

1. If the study was to be repeated the present items would need to be analysed and altered or discarded as indicated.

2. The Somewhat False and Mostly False categories in the questionnaire are not entirely satisfactory since items in these categories must also be Mostly True and Somewhat True respectively. However, when questioned on this point only one of the fifty-six students responding to the instrument had considered this possibility.

3. More serious is the failure to pre-decide the conditions under which an individual's responses would be considered invalid and discarded. For example, one of the experimental subjects responded to 33 of the 45 items as Entirely False and the other 12 as Entirely True. Both the Chi-square and the Kolmogorov-Smirnov One-sample test indicate the rejection of the hypothesis that such a score could be made by a member of the experimental sample. The elimination of this single paper would have made appreciable differences in the outcome of the experimental findings.

### REFERENCES

1. Storey, Arthur G. "Changes in Attitude and Behavior as a Function of Methods Courses," *The Alberta Journal of Educational Research*, Vol IX, No. 1, March, 1963.
2. Chorny, M. and Storey, A. G., "A Team Approach to Practice Teaching," Unpublished at present.
3. Siegel, M., *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Company, Inc., New York, 1956, pg. 47.



# AN EXPERIMENTAL APPLICATION OF THE PRINCIPLE OF INCONGRUITY TOLERANCE TO THE COUNSELLING SETTING

G. B. JONES AND C. M. CHRISTENSEN  
*University of Alberta, Edmonton*

## Problem

Recent literature in dynamic psychology has involved the postulation of various hypothetical constructs which apparently are related. One group of these concepts centers around the behavioral phenomenon experienced by individuals encountering seemingly opposing cognitive elements (knowledge, beliefs, attitudes or opinions). Such terms as the following have been used by various writers to refer to this phenomenon: ambiguity, inconsistency, uncertainty, conflict, incompatibility, non-conformity, imbalance, asymmetry, dissonance, incongruity and disharmony.

Rokeach (1960) has stated that it would be

a worthwhile task to examine more closely the extent to which these various concepts overlap each other, the respective ranges of their application, and how they may be more integrated with each other (p. 395).

This "task" was attempted in a fairly extensive review of the literature and because of space limitations will not be reported here other than to note the concepts advanced by a few of the more eminent theorists in the area. From this review, however, an experiment studied the feasibility of placing subjects in a learning situation aimed at promoting "incongruity tolerance". This paper describes the experimental study, its results and its implications for teachers and counselors.

## Theoretical Framework

In many cases the aforementioned terms have been used to identify the same concept which describes the existence of conflict in the social, physical or mental environment of the individual. This conflict usually induces tension or emotional arousal causing the individual to take immediate steps to resolve the resultant unpleasant state. It would appear that the following paradigm of incongruity applies to all of the related concepts identified by these diverse terms:

incongruity → tension → behavior → congruity

This is a paradigm of operant conditioning; habitual behavior for most individuals.

In this paper then the term "incongruity" refers to the resultant state experienced by the individual when two or more unfitting or dissimilar cognitive elements interact.

The theoretical exposition most fundamental to this study is that of Festinger (1957) who expounded the concept of "cognitive dissonance". Considering pairs of cognitive elements as basic to behavior or knowledge, Festinger considered that dissonance arises when the obverse of one cognitive element follows from the existence of the other element of that pair.

The existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance. (p. 6)

This approach, that the existence of dissonance (or incongruity) will almost automatically lead to a striving for congruity, is similar to that taken by Osgood and Tannenbaum. They use the term "congruity" in relation to the connotational meaning of concepts as measured by Osgood's semantic differential scale. It describes an harmonious unity between an individual's overt attitude to or evaluation of a concept and his judgmental frame of reference. They also use the term to describe the relationship between an individual's evaluative attitude toward a source of information and his attitude toward the concept imparted by this source.

Not too dissimilar but used more in the direction of the act of perception is Bruner's approach (1949). Perceptual incongruity is said to occur when an individual's set or expectancy regarding perception of stimuli is not confirmed. Such incongruity can occur as a result of

an unexpected concatenation of events, a conspicuous mismatching or an unlikely pairing of cause and effect. (p. 08)

In addition, Taguiri, Bruner and Blake (1958) introduced the term perceptual-affective congruency to describe the extent to which an individual sees the feelings of others for him as corresponding to his feelings for them.

Berlyne (1960) dichotomizes perceptual and conceptual incongruity as two of the aspects of conceptual conflict. Conceptual incongruity, involving symbolic responses, arises when an individual finds that something he has formerly believed authentic and regular now has an exception. Such a concept is not too different from Festinger's "cognitive dissonance". Berlyne defines perceptual conflict, involving perceived stimuli, as arising when stimuli which formerly occasioned certain expectatons on the part of the perceiving organism are now perceived in conjunction with stimuli which disappoint these expectations. A distinct resemblance to statements of Bruner and Postman (1947).

The structure of belief-disbelief system is studied by Rokeach (1960) who reviews this structure in relationship to people's open and closedmindedness. He postulates that the more open one's belief system then the more should incoming information be evaluated on its own merits, the more should the individual be governed by internal self-actualizing forces and the more should he be able to resist externally imposed reinforcement. This definition of openness is considered in this paper as being descriptive of the individual who has approached a level of maximum tolerance of incongruity.

Rokeach also refers to the concept of "belief congruence" noting that

each person is somehow motivated to arrange the world of ideas, of people, and of authority in harmonious relations with each other. (p. 395)

Further description of maximum tolerance of incongruity is added by Harvey, Hunt and Schroder (1961) who theorize on the structural characteristics of conceptual systems. They have developed a personality theory based on the stage and level of cognitive functioning of the individual. Progression is made from the concretistic or absolutistic stage to the fourth and final stage of development structured with internal and informationally based concepts which have maximum abstractness. An individual at this level of cognitive functioning (as well as one at a level of maximum tolerance of incongruity) would have optimal tolerance of stress or threat, being less likely to resort to defensive resolutions of anxiety than someone at a lower stage. Their statement that the basis for progressive conceptual development to occur is the ability to tolerate refutation of conceptual systems is quite fundamental to the theoretical framework of this present study.

The concept of the interpersonal matrix was introduced by Secord and Backman (1961) to explain personality and behavioral stability and change. This matrix is a composite of three interacting components: an aspect of the individual's self-concept, related interpretations of his own behavior, and his perception of the related aspects of the behavior of the individual with whom he is interacting. Here the individual strives to maintain congruency among these components by shaping the interaction process through manipulation, control or distortion of one or two of the components.

In somewhat a similar approach Heider (1946) spoke of the configuration or unit relationship of person-person, person-nonperson, two persons-nonperson, or three persons as being the fundamental concept in behavioral dynamics. He used the term "symmetrical harmony" to describe the existence of a "balanced configuration" where all the entities have a two way either positive or negative attitude toward each other. These "entities" can include persons,



situations, events, ideas or things. If balance does not exist then forces toward this state will arise and if change is impossible, then tension will result.

It appears that the concepts reviewed do overlap significantly and that the change in terminology by the theorists has paved the way for much repetition of ideas. The existence of congruent states toward which incongruent individuals strive does seem to be accepted as factual by most writers. Complete support also is given to the idea that the existence of incongruity usually produces behavior aimed at its reduction but few writers explained the motivational dynamics of this phenomenon. No mention was made concerning whether this striving is autochthonous and universal or whether it is learned and dependent upon individual differences.

Most of the approaches stressed the noxious aspects of incongruity, apparently evaluating congruous states as valued objectives. Berlyne (1960) and Harvey, Hunt and Schroder (1961) were exceptions. The former pointed out that all aspects of conceptual conflict provide

the motivation for intrinsic epistemic behavior, and supplement extrinsic motivation when knowledge is sought for practical or social ends. (p. 280)

Such statements indicate the role that can be played by an ability to tolerate incongruity.

Very little mention was made of individual differences in tolerance of incongruity or of the possibility that such tolerance could be profitable for the incongruent individual. Most of the concepts reviewed seemed to regard the process of striving for congruity as a natural outcome of the existence of incongruity: a process initiated by fortuitous circumstances which disappoint the individual's expectations.

### **Postulates From Related Theory**

1. There are individual differences in ability to tolerate incongruity. However, most individuals in Western cultures habitually seek congruous states, desiring to avoid incongruity as easily and as quickly as possible.
2. Both tolerance and intolerance of incongruity are learned personality traits and as such can be changed.
3. If the positive incentives are recognized, an individual can learn to be tolerant of incongruity.
  - (a) There will still be an induction of arousal when incongruity is encountered but defensive reactions will not be evoked immediately and perhaps not at all.

(b) An individual, recognizing the advantages of tolerating the incongruity for a period of time, will be able to control and channel the arousal. Progress could be made to a level where control is not necessary and the individual could live with the arousal without tense or restricted feelings. Ultimately the individual could find incongruity quite pleasurable.

(c) Positive incentives for incongruity tolerance could include: realization of a positive type of freedom—"freedom to"—including more inner-directedness (self control) and greater pursuit of personal interests; more openmindedness to novel stimuli and to potentially refuting stimuli; progression toward more abstract stages of conceptual functioning; more creativity and spontaneity; more original and critical thought; and increased epistemic behavior.

### **Hypothesis**

Provision of a learning situation in which the positive incentives of incongruity tolerance are expounded will result in personality changes predictable from the theoretical framework postulated here.

### **Procedure**

#### *Subjects*

Forty-one volunteers (twenty-five females) from nine tutorial classes in an introductory course in Educational Psychology at the University of Alberta comprised the sample for this study. The subjects were randomly assigned to one experimental and two control groups. The mean age for the group was 20.6 with a standard deviation of 3.6.

#### *Method*

After the first session of testing, two groups met separately for one hour a week for the next seven weeks for informal discussion sessions while the other control group returned only on the ninth week for retesting. The experimental group was led in discussions on the theory behind tolerance of incongruity during the weekly sessions. Attempts were made to have the subjects rephrase the concepts in their own words and provide personal examples of them. Refutation of the concepts provided by the leader was particularly encouraged. Most of each session was assigned to application of the advocated concepts to reality and the subjects were encouraged to think of further personal applications between sessions.

The weekly sessions of the one control group were conducted by the same leader and here discussion centered around topics of personal interest. Most of the discussions stemmed from personal problems or from Educational Psychology course work. All subjects were requested not to discuss group progress with members of another group for the duration of the experiment.



During the retesting session members of the experimental group were also administered an open-ended questionnaire to ascertain their opinions on the topic of incongruity tolerance and intolerance, its personal and educational significance, the effectiveness of presentation of the topic, and personal and interpersonal changes noted as a result of the group discussions.

### *Testing Instruments*

Tests specifically designed to measure personality changes which would result from an increase in incongruity tolerance have not been designed. Therefore a variety of personality instruments was chosen and an attempt was made to compare changes measured by these instruments to those predictable from the theory underlying the concept of congruity.

*Rokeach's Dogmatism Scale.* This test was constructed to measure individual differences in open and closed belief systems and general authoritarianism and intolerance although no validation information has been obtained regarding its success with the latter two variables. Rokeach states that on the Scale high scorers showed less ability to receive, evaluate and act on relevant information and less ability to form new belief systems. It was felt that subjects who increased in incongruity tolerance would make lower scores on this Scale.

The Dogmatism Scale consists of forty items designed to measure the structure rather than the content of belief systems. Most validity studies have not been too successful. Reliability coefficients ranging from .68 to .93 and clustering around .80 are given for samples of American and English college students, English workers and American veterans.

*Edward Personal Preference Schedule.* This instrument provides measures on fifteen personality variables based on two hundred and twenty-five pairs of items. Subjects must indicate the statement in each pair which is most characteristic of their likes and feelings. An attempt has been made to minimize the influence of social desirability on scale values.

Norms (based on normative samples of United States college students) are provided for both sexes for each personality variable and a consistency variable. Quite extensive supportive data has also been accumulated including: coefficients of internal consistency, stability coefficients, intercorrelation of the test's personality variables and validity coefficients gained from comparisons made with the Guilford-Martin Personal Inventory and the Taylor Manifest Anxiety Scale.



*Texture Preference Inventory.* This test, developed by Dr. Clifford M. Christensen of the University of Alberta ,is composed of one hundred and forty-five pairs of photographs of textures. The method of presentation is similar to that used in Edwards P.P.S. This experimental test, based on the hypothesis that preference for certain types of textures is functionally related to specific personality variables, gives scores on twelve scales. These scales can be grouped into three types of preferences including: preference for complex, ambiguous or unstructured textures; preference for simple ambiguous textures which could be described as dull and uninteresting; and preference for simple structured textures.

This test is of recent inception and thus, although reliability and validity data have been collected, none has as yet been published. The test was used in a secondary role to compare its results to those attained by the tests of Edwards and Rokeach. It was postulated that subjects who increased in tolerance of incongruity would show an increase in preference for ambiguous textures.

Results

An analysis of variance for a one-way classification, as in Ferguson (1959), was computed to examine the differences between pre-treatment and post-treatment scores for all groups on all the variables of the three personality tests. The five and one percent levels of significance were accepted as the desired levels and thus F ratios equal to or greater than 3.25 and 5.21 were required for significance at these levels respectively. The groups are designated by letters as follows: Group "A" is the experimental group, Group "B" is the control group which met in weekly discussional sessions, while Group "C" is the non-attendance control group.

TABLE I  
ANALYSIS OF DIFFERENCES BETWEEN PRE- AND POST-TREATMENT SCORES FOR DOGMATISM SCALE

Personality Variable	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\bar{d}$	$\sum d^2$	$\bar{d}$	$\sum d^2$	$\bar{d}$	$\sum d^2$		
Dogmatism .....	1.57	4,390	3.69	4,732	-1.00	3,192	.24	ns

As shown by Table I changes indicated by Rokeach's Dogmatism Scale were far from significant primarily because of large within-group variances.

Table II gives the results compiled on the Edwards Personal Preference Schedule. Significant changes were found for the personal-

ity variables of Affiliation (.005 level), Autonomy (.01 level) and Endurance (.025 level). After treatment the subjects of the experimental Group A scored lower on the Affiliation variable while the control group subjects scored higher over the same period of time. Group A subjects showed greater increases on the Autonomy scale than did the subjects of Group B while those in Group C decreased on this variable. The Endurance variable resulted in no change for Group A subjects while the other groups' subjects attained lower scores on this scale.

Although the findings were not significant, the subjects in Group A did tend to attain lower scores on the Succorance variable and higher scores on the Dominance variable than did the subjects in the control groups.

TABLE II  
ANALYSIS OF DIFFERENCES BETWEEN PRE- AND  
POST-TREATMENT SCORES FOR EDWARDS PERSONAL  
PREFERENCE SCHEDULE

Personality Variables	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\bar{p}$	$\sum d^2$	$\bar{p}$	$\sum d^2$	$\bar{p}$	$\sum d^2$		
Achievement .....	.79	165	— .15	106	1.0	130	.49	ns
Deference .....	.07	99	—1.62	167	—1.57	84	1.72	ns
Order .....	—1.79	225	—1.87	196	— .21	149	.94	ns
Exhibition .....	2.21	227	.45	112	1.79	183	.93	ns
Autonomy .....	2.50	245	1.31	117	—1.07	75	5.64	p < .01
Affiliation .....	—2.36	222	2.23	187	1.50	193	7.41	p < .005
Intracception .....	— .57	102	—1.23	218	.64	61	1.33	ns
Succorance .....	— .93	233	.92	198	0	254	.66	ns
Dominance .....	2.57	428	.54	161	— .21	119	1.80	ns
Abasement .....	— .57	268	.51	106	—1.56	172	.69	ns
Nurturance .....	—2.64	241	0	154	—1.92	233	1.99	ns
Change .....	.36	109	— .46	136	.86	234	.49	ns
Endurance .....	.21	119	—3.08	256	— .21	111	4.48	p < .025
Heterosexuality .....	1.28	172	1.92	321	.71	206	.30	ns
Aggression .....	— .71	148	.92	86	—1.50	159	2.27	ns

TABLE III  
ANALYSIS OF DIFFERENCES BETWEEN PRE- AND  
POST-TREATMENT SCORES FOR TEXTURE  
PREFERENCE INVENTORY

Variables	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\bar{p}$	$\sum d^2$	$\bar{p}$	$\sum d^2$	$\bar{p}$	$\sum d^2$		
I .....	.29	192	— .23	73	— .71	124	.35	ns
II .....	2.57	360	1.77	159	1.57	98	.33	ns
III .....	—1.71	292	1.00	73	—1.57	98	3.16	$p < .06$
IV .....	—1.07	191	— .15	76	—1.69	123	1.04	ns
V .....	1.28	90	— .23	71	— .14	128	1.43	ns
VI .....	2.21	225	— .92	136	1.36	189	3.01	$p < .10$
VII .....	.07	143	— .46	132	— .21	117	.12	ns
VIII .....	1.14	172	— .54	83	— .21	127	1.03	ns
IX .....	—2.29	324	— .54	217	— .43	176	.89	ns
X .....	—2.36	377	— .77	102	— .07	105	1.46	ns
XI .....	2.29	340	0	82	—1.14	116	3.63	$p < .05$
XII .....	—2.00	324	— .51	192	4.14	572	6.77	$p < .005$

NOTE: Variables I, II, V, VIII, XI contain complex ambiguous (unstructured) textures.  
Variables III, IV, IX contain simple structured textures.  
Variables VII, X, XII contain simple ambiguous (unstructured) textures.  
Variable VI contains unstructured textures.

In Table III, significant differences among group scores can be noted on variables XI (.05 level) and XII (.005 level) of the Texture Preference Inventory. After treatment the subjects of Group A tended to increase in preference for the complex-ambiguous textures of variable XI and to decrease in preference for the simple unstructured textures of variable XII. The subjects of Group C showed opposite tendencies while those of Group B remained somewhat the same on both variables.

Once again non-significant trends were noted on a number of the variables of this test including: I, II, V, VI, VIII, IX and X. These changes were in the direction predicted by the basic concepts tested here, for the subjects of Group A consistently tended to



show greater increases in preference for the complex ambiguous textures. They also tended to display a greater decrease in preference for the simple structured and simple abstract textures of the test.

### Discussion

It is felt that the findings on the Personal Preference Schedule and the Texture Preference Inventory support the hypothesis that significant personality changes in line with incongruity tolerance theory would result after treatment. The Affiliation variable of Edwards' test purportedly is associated with manifest needs for participation in groups, for formation of friendships and for action in groups rather than individual action. Therefore it is suggested that because the individuals in Group A had learned to tolerate incongruity more adequately and now needed less social support, they scored significantly lower on this variable. These subjects increased significantly on the Autonomy variable and because it is contended that this variable is related to such needs as those for independence in decision making, freedom from convention and conformity, and indulgence in critical thought and action, it is further suggested that additional support is given to the premise that the group increased in tolerance of incongruity.

Further support for the preceding statement was gained from the Texture Preference Inventory. Group A did tend to increase in preference for the ambiguous textures and to decrease in preference for the structured textures. This test was constructed on the assumption that subjects who consistently exhibited preference for clusters of complex-ambiguous or unstructured textures would have fairly high ambiguity tolerance, a willingness to accept a state of affairs which is subject to alternate interpretations or outcomes.

Rokeach's Dogmatism Scale showed no significant changes in personality variables among the groups and either a discrepancy in this present study or in the construct validity of this test could be inferred. It was thought that if the subjects of Group A increased in tolerance of incongruity, they would display more open-mindedness. Rokeach implied that persons with open belief systems would score lower on his Scale (and thus be less dogmatic). The results of this study only showed a tendency for the groups which met in weekly discussions to make higher scores after the treatment period. The possibility that many persons could be classified as openminded and that they could support their belief systems quite dogmatically appears to have been eliminated from Rokeach's theoretical framework. It could therefore be suggested that Rokeach's instrument attempts to measure the personality variables of dominance and autonomy and that it does not provide

too reliable a measure of the extent to which individuals are dogmatically closed to new belief systems or the degree to which they are open.

Further experimentation involving this concept of incongruity tolerance would seem necessary. Such experimentation must involve a setting in which the effects of interpersonal, intrapersonal and cultural or normative forces can be studied. An attempt should also be made to use a more representative sample of subjects than was used in the present study. If subjects from cultures other than the Western culture were used it could possibly be determined more positively whether intolerance of incongruity is a universally inherited behavioral trait or one which is acquired from interaction in certain environmental settings. A followup study would also be advisable to determine the retention of incongruity tolerance over a longer period of time. All further experimental attempts would depend upon the development and documentation of an instrument which can provide a valid estimate of the presence and extent of an ability to tolerate incongruity.

The motivational components of incongruity intolerance should also be studied by experimental methods. Perhaps it would be found that the basic motive behind an habitual striving for congruous states is the desire to avoid the feelings of uncertainty and insecurity arising from not knowing or understanding a state of affairs, especially when it is important for the security of the individual involved or when it runs contrary to his expectations.

It could be that tolerance of incongruity is the fundamental principle of all successful psychotherapies and involves nothing more than learning to live with tension, learning to recognize incompatibilities and accept them without defensive reactions. It could be postulated also that incongruity tolerance is a function of intelligence and thus persons at lower levels are not capable of tolerating incongruous states.

The present investigation did not seem to support Rokeach's (1960) apparent combining of the traits of dogmatism and closed-mindedness. Further investigation of this finding could be profitable. Harvey, Hunt and Schroder (1961) seem to disagree with Rokeach also. They claim that the more abstract the conceptual system, the greater the resistance within that system to increasing closedness. They state that the more abstract levels of functioning are characterized by

the ability to hold strong beliefs and to be ego-involved, while remaining open to alternate evaluations and differentiations, that is, to hold strong beliefs but to be unprejudiced. Such functioning permits ego-involvement without the usual accompanying disadvantages of bias, subjectivity, and distortion. (p. 195)



In educational and psychotherapeutic fields openness of thought and progression toward tolerance of incongruity should be encouraged. In children such progression could be impeded by the encouraging of rigid identification with parental and authoritative models (e.g. teachers), by the absolute presentation of materials and facts with no opportunity allowed for refutation, by requiring that students learn only that which is accepted by the instructor, or by presenting examinations that involve absolute regurgitation of material expounded by the instructor.

Finally it could be implied that incongruity tolerance is positively correlated with creativity, independent and critical thought, and flexibility of personality. It would be further suggested that conceptual structures characterized by such traits are the desirable goals toward which educational and guidance services should encourage students to progress. The basic requirement in providing environmental opportunities for the cultivation of such traits would be that educators and counselors first possess these traits themselves.

#### REFERENCES

1. Berlyne, D. E., *Conflict, Arousal and Curiosity*, New York, McGraw-Hill Co., 1960.
2. Bruner, J. S., Goodnow, J. J., Austin, G. A., *A Study of Thinking*, New York, John Wiley and Sons Inc., 1956.
3. Bruner J. S., Postman, L., "On the Perception of Incongruity: A Paradigm", *Journal of Personality*, 1949, 18, 206-223.
4. Ferguson, G. A., *Statistical Analysis in Psychology and Education*, New York, McGraw-Hill Co., 1959.
5. Festinger, L., *Cognitive Dissonance*, Evanston, Illinois, Row, Peterson and Co., 1957.
6. Harvey, O. J., Hunt, D. E., Schroder, H. M., *Conceptual Systems and Personality Organization*, New York, John Wiley & Sons, 1961.
7. Heider F., "Attitudes and Cognitive Organization", *The Journal of Psychology*, 1946, 21, 107-112.
8. Heider, F., *The Psychology of Interpersonal Relations*, New York, John Wiley and Sons, Inc., 1958.
9. Osgood, C. E., Tannenbaum, P. H., "The Principle of Congruity in the Prediction of Attitude Change", *Psychological Review*, 1955, 62, 42-55.
10. Rokeach, Milton, *The Open and Closed Mind*, New York, Basic Books, Inc., 1960.
11. Secord, P. F., Backman, C. W., "Personality Theory and the Stability and Change in Individual Behavior, an Interpersonal Approach", *Psychological Review*, 1961, 68, 21-32.
12. Taguiri, R., Bruner, J. S., Blake, R. R., "On the Relation Between Feelings and Perceptions of Feelings Among Members of Small Groups", in *Readings in Social Psychology*, S. E. Maccoby, T. M. Newcomb, E. L. Hartley, New York, Henry Holt & Co., 1958, 110-116.



# STEREOTYPING AND THE USE OF ALL-INCLUSIVE TERMS

JAMES G. SNIDER

*University of Alberta, Calgary*

The process of stereotyping has been extensively investigated. Lippmann (1922, p. 88) very early pointed out the importance of stereotypes. He said that stereotypes were the pictures we carry in our heads: "We do not so much see this man and that sunset; rather we notice that the thing is man or sunset, and then see chiefly what our mind is already full of on those subjects." And "foreigners," "Baptists," and "Communists" are similar in this sense to men and sunsets. The reviews of Klineberg (1950), Harding, *et al* (1954), and Vinacke (1956) show that stereotyping has many correlates. Stereotyping has been shown to influence ratings of personality traits (Razran, 1950), judgments of political slogans (Lewis 1947), and the identification of group membership from physical appearance (Lund and Berg, 1946). On the other hand, some variables influence the individual's tendency to stereotype. Stautland (1959) found intelligence to be negatively correlated with "strength" of stereotyping. The likelihood of a group being stereotyped may be affected by the actions of that group (Meenes, 1943). Stereotyping may also be affected by the tendency of the individual to use projection as a defense mechanism (Child and Doob, 1943). It is the aim of the present study to delineate still another variable which may influence the individual's tendency to stereotype. The idea for the study arose from the observation that stereotypes are often stated in actual or implied all-inclusive terms. "Schoolboys are lazy"—implying **all** are lazy. "Professors are absent-minded." **All** of them? **Always**? In **every** circumstance? The general question of the study concerned whether or not those who tend toward an excessive use of all-inclusive terms, such as **all**, **always** and **never**, will not also tend to stereotype more than usual.

## Relevance of the Study for Education

It has often been pointed out that it is important in inter-personal relations to guard against stereotyping and important in education to teach against stereotyping (e.g., Klineberg, 1950), and further it has been said that to do these things effectively we need to know as many of the correlates of stereotyping as possible. This study is an attempt to examine for yet one more variable which might help us guard against and teach against stereotyping.

### Hypothesis

It was hypothesized that individuals who tend to a more than average use of such all-inclusive terms as **all**, **never**, and **always**, will have a greater tendency to stereotype than will individuals who tend to a less than average use of such terms. The problem of the study, therefore, was stated as follows: Given two groups, one with a greater tendency than the other to use all-inclusive terms will have a greater tendency to stereotype than will the group with the lesser tendency to use all-inclusive terms.

### Design of the Study

In selecting a population for the study, individuals of approximately similar socio-economic condition, educational level, and age were sought. Accordingly, the two instruments used were administered to 292 randomly chosen tenth-grade pupils in a large urban high school in Calgary, Alberta.

An instrument devised by the present writer was used as a measure of the tendency to use all-inclusive terms. The test consists of 80 items, 50 "crucial" items and 30 "distractor" items, introduced to insure against pupils discovering the purpose of the test from simply scanning the items, as might be possible, since the items are simple in design. Interviews showed that the distractor items did indeed seem to serve their purpose: none of more than thirty pupils interviewed seemed to have any notion of the true purpose of the test. They were willing to accept at face value that the test was simply to find some "word preferences" of pupils their age.

The "crucial" items of the test were constructed from pairs of words chosen so as to be as nearly alike as possible and yet so that one word of each pair is an all-inclusive term and the other is not and so that the words in each pair were similar in frequency as indicated by Thorndike and Lorge (1944). It is assumed that in such a situation the individual who tends to use all-inclusive terms will tend to prefer the all-inclusive term of each pair. An item-analysis was carried out, utilizing the biserial correlation coefficient, and only items with a coefficient significant at the 5 per cent level of confidence were retained to form the test.

The test has adequate reliability for testing groups: several Kuder-Richardson and split-half reliability coefficients approximate .90. Two validity coefficients were obtained. In a group of 60 ninth-grade pupils, a rank correlation of .47 was obtained between total score and word counts of all-inclusive terms used in free writing. For a group of 48 tenth-grade pupils, the correlation was .36.

The directions for taking the test are simply to choose the preferred word of each pair. The score is the number of all-inclusive terms chosen. The first 20 items of the test, with the all-inclusive terms in the "crucial" items underlined are as follows:

1. wander	stray	11. miserable	unhappy
2. loosen	unfasten	12. <b>entirely</b>	largely
3. seldom	<b>never</b>	13. <b>always</b>	often
4. detect	reveal	14. peril	danger
5. <b>are</b>	seem	15. opinion	knowledge
6. poverty	desolation	16. worst	bad
7. destiny	fate	17. <b>must</b>	can
8. <b>certain</b>	likely	18. assert	declare
9. now	soon	19. maybe	<b>surely</b>
10. den	cave	20. <b>completely</b>	mostly

As a measure of stereotyping, a test devised by Stautland (1959, p. 14) was used. This test, constructed after the manner of Osgood's (1957) semantic differential, places the respondent in the situation of judging various concepts: Englishmen, Chinese, Americans, Mexicans, Russians, Norwegians, in regard to pairs of polar terms: practical-impractical, cruel-kind, intelligent-dumb, inferior-superior, happy-sad, dirty - clean, brave - cowardly, warlike - peaceloving, honest-dishonest, lazy-hardworking. The respondent judges in terms of a seven-point scale, weighted as shown:

Polar Term	(3	(2	(1	(0	(1	(2	(3	Polar Term
	<u>  </u>	<u>  </u>	<u>  </u>	<u>  </u>	<u>  </u>	<u>  </u>	<u>  </u>	

Scores on the test may range from 0 to 180, and it is assumed that the higher the score the greater the tendency to stereotype. Stautland obtained test-retest reliability coefficients over .90 with his test and a validity coefficient of .39 with a test of stereotyping devised by Siegel (1954).

In terms of total score on the test of the tendency to use all-inclusive terms, the sample was divided into the upper and lower 25 per cent, giving us two groups, one with a stronger tendency than average to use all-inclusive terms, the other with a lesser tendency than average to use such terms. These two groups were then compared as to their scores on the stereotyping test.

### Results

Table I presents a summary of the results of the study. It seems quite clear that the differences shown in the table are not likely to be due to chance alone, as the tests for differences in means indicate. It would appear that the prediction for the study was corroborated. The relationship between the tendency to use all-inclusive terms and stereotyping is apparently a real one and can be empirically demonstrated.



TABLE I  
MEANS FOR THE UPPER AND LOWER 25% ON THE  
TEST OF THE TENDENCY TO USE ALL-INCLUSIVE  
TERMS AND MEANS FOR SCORES ON THE  
STEREOTYPING TEST

Test of Tendency to use All-Inclusive Terms	Means		Stereotyping Test Means	
Upper 25% .....	37.2	$P < .001$	100.7	$P < .01$
Lower 25% .....	14.4		90.2	

The general tendency of the results shown in the statistical presentation was also observed in the examination of individual cases when pupils had received very high or very low scores on the test of the tendency to use all-inclusive terms. One boy, for example, who obtained the highest score, 49, on the test of the tendency to use all-inclusive terms, also obtained the highest score, 180, on the stereotyping test. The opposite could be seen in pupils who scored very low on the test of the tendency to use all-inclusive terms.

### Discussion

While the results of the study suggest that the use of all-inclusive terms is a correlate of stereotyping, it is by no means clear whether or not there is any causal relationship between the two variables. Both the use of all-inclusive terms and the tendency to stereotype may be merely symptoms of a more general "tendency to be all-inclusive." It may be, however, that the use of all-inclusive terms is a contributing factor in stereotyping. More experimental work needs to be done to answer such a question, but the results of the present study suggest that such an undertaking might be worthwhile. Whatever the causal connection between the variables of the study, teachers and others concerned with stereotyping can assume that where there is an excessive use of all-inclusive terms there is also likely to be a tendency to stereotype. If we assume that the use of all-inclusive terms is a contributing factor to stereotyping and that a lessening of the use of all-inclusive terms might curb the tendency to stereotype, we need some means of analyzing the person who uses all-inclusive terms to excess. Fortunately, various therapeutic suggestions (Hayakawa, 1943; Johnson, 1946; Korzybski, 1948), which might be put to use by teachers and others, are available already for dealing with the tendency to an excessive use of all-inclusive terms.

## REFERENCES

- Child, I. L., and Doob, L. W. Factors Determining National Stereotypes, *J. Soc. Psychol.*, 1943, 17, 203-219.
- Harding, J., Kutner, B., Proshansky, H. and Chein, I. Prejudice and Ethnic Relations. In *Handbook of Social Psychology*, Vol. II, Gardner Lindzey, Ed. Reading, Mass.: Addison-Wesley, 1954.
- Hayakawa, S. I. *Language in Action*. New York: Harcourt, Brace, 1943.
- Johnson, W. *People in Quandries*. New York: Harpers, 1946
- Korzybski, A. *Science and Sanity*. Lakeville, Conn.: The Institute of General Semantics, 1948.
- Klineberg, O. Tensions Affecting International Understanding: A Survey of Research, *Soc. Sci. Res. Coun. Bull.*, 1950, No. 62.
- Lewis, H. B. An Experiment on the Operation of Prestige Suggestion, in T. M. Newcomb and E. L. Hartley (Eds.) *Readings in Social Psychology*. New York: Holt, 1947, pp. 232-243.
- Lippmann, W. *Public Opinion*. New York: Harcourt, Brace, 1922.
- Lund, F. H. and Berg, W. C. Identifiability of Nationality Characteristics, *J. Soc. Psychol.*, 1946, 17, 327-336.
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. *The Measurement of Meaning*. Urbana, Illinois; U. of Illinois Press, 1957.
- Razran, G. Ethnic Dislikes and Stereotypes: A Laboratory Study. *J. Abnorm. Soc. Psychol.*, 1950, 45, 7-27.
- Siegel, S. Certain Determinants and Correlates of Authoritarianism. *Genet. Psychol. Monogr.*, 1954, 49, 187-229.
- Stautland, S. Some Correlates and Determinants of National Stereotyping. Unpublished doctoral dissertation, Stanford University, 1959.
- Thorndike, E. L. and Lorge, I. *The Teacher's Word Book of 30,000 Words*. Columbia University Press, 1944.
- Vinacke, W. E. *The Psychology of Thinking*. New York: McGraw-Hill, 1952.
- Vinacke, W. E. Explorations in the Dynamic Processes of Stereotyping. *J. of Soc. Psychol.*, 1956, 43, 105-32.

## BOOK REVIEW

### TEAM TEACHING

*Authors:* Judson T. Shaplin and Henry F. Olds, Jr. (eds.) and others.

*Publishers:* Harper and Row, New York, 1964

In the foreword, Frances Keppel asks for a detailed assessment of team teaching and says that "the time has come to sharpen definitions, to puncture some balloons and to put the movement into proper perspective." Most people will agree and this publication makes a commendable effort to accomplish just this.

Written by a number of authors it suffers from some unevenness of style and of subject matter. However, its major strength lies in a clearer definition of team teaching and a development of the theoretical bases on which various aspects of team teaching rely. For example the team concept is related to social organization, small group and organizational theory. Another section developed a useful taxonomy of teams in terms of functional variables.

Educators will be pleased to see also a detailed section on the type of research needed at different stages of team teaching development. Useful information of a different order is given on historical antecedents of team teaching, curriculum adaptation, relationship with the educational reform movement, school buildings and details of many current projects.

Canadian administrators interested in this major innovation of today, as well as the many sitting on the fence waiting, should welcome this appraisal, the main authors of which, though deeply involved in team teaching since its inception six years ago, have allowed their bias to show but rarely.

W. D. Neal

### BOOKS RECEIVED

*The Scottish Scholastic Survey*, 1953. University of London Press Limited, 1963. Pp. 211. 25s

A total of approximately 72,400 children ranging in age from 9 years 11 months to 10 years 10 months were given a battery of tests in English and arithmetic. The tests used, keys, norms and variances are presented in detail. A chapter of conclusions and recommendations follows.

Dobinson, C. H., *Schooling, 1963-1970*. George G. Harrap, London and Toronto, 1963. Pp. 175. \$3.15.



A critical survey of the British educational system, suggesting changes to meet the needs of changed conditions of life. Discontents are continually expressed with various aspects and levels of the system and with the studies themselves, and suggestions for modification are put forward from many quarters.

Although the author's hard-hitting remarks apply primarily to British education, Canadian educators will find them pertinent to the local scene.

Gardner, D. Bruce, *Development in Early Childhood: The Preschool Years*. Harper and Row, New York. 1964. Pp. 349. \$5.75.

A solid introduction to the field, this text organizes the factual information on the child's first six years around a unifying theme of self-development. While recognizing the significance of physical growth, the book gives relatively greater emphasis to achievement of self-hood.

A brief overview prefaces each of the four parts of the text—How We Study Children, Foundations of Development, Aspects of Development in the Preschool Years, and Society of the Pre-School Child.

Gordon, Ira J. *Human Development from Birth through Adolescence*. Harper and Brothers, N.Y. 1962. Pp. 400. \$5.75.

This text considers in thorough fashion the development of a newborn infant into an adult. It combines an external approach and an internal approach to the study of the growth of human beings. It will prove useful to those studying and aiding infants in their development into productive adults. The main purpose of the book is to present theoretical and other information which treats the significant relationships among human development in general, general behavior, and the development of the self in each individual.

The Editorial Committee for the Alberta Journal of Educational Research will be pleased to consider research articles from contributors outside the Province of Alberta, for publication in the Journal.

Manuscripts should be addressed to the Editor.

H. E. Smith





# The Alberta Journal of Educational Research

Vol. X, No. 4

December, 1964



THE COMMITTEE ON EDUCATIONAL RESEARCH  
*Faculty of Education*  
*University of Alberta*



# SIR RICHARD LIVINGSTONE ON EDUCATION

BARBARA HUTCHINSON

*University of Alberta, Calgary*

The purpose of this study is to examine the educational philosophy of Sir Richard Winn Livingstone, a contemporary English educator, and to relate it to current theory and practice. His ideas not only contribute to the history of educational theory generally but also offer us constructive suggestions to make our educational practices more appropriate to the needs of the twentieth century.

## **Livingstone's Interpretation of Ideological Changes in the Contemporary World**

To appreciate the contributions of Sir Richard Livingstone (and hence to appreciate contemporary educational problems), one must examine the milieu in which he lived and worked—"the character of his age, the influences social, intellectual and other which have shaped his outlook and his intellect."<sup>1</sup> He regarded the twentieth century as a period of moral confusion similar to that of fifth-century Athens. Such forces as liberalism, rationalism and science not only have changed the material world and society but also have altered "the very texture of our thought", their impact striking violently on traditional beliefs and throwing them into confusion. Science has brought about the annihilation of space which made a new international outlook necessary. Unfortunately, this broader outlook has not been forthcoming; anachronistic isolationism and provincialism are still prevalent. Science has given us the means to abolish poverty, but we are too uneducated to spend our wealth intelligently. Science has provided us with fine material possessions, but we have not found a substitute for the satisfaction of craftsmanship. Science has provided us with analytical techniques upon which the maintenance and progress of our civilization depend, but we have tended to mistake an analysis of things for a knowledge and appreciation of them. We must not forget the dual nature of our civilization, compounded as it is of the material world and the human being. The great obstacles to individual, national and international progress arise from the latter, from human conduct. It was to the solution of these problems through education that Sir Richard Livingstone addressed himself.

### **Influences on Livingstone**

Three factors had a profound influence upon Livingstone: Hellenism, Christianity and the Danish-People's High School movement.

1. Sir Richard Livingstone. *Education and the Spirit of the Age*. Clarendon Press, Oxford, 1952, p. 80.



In him these influences "formed a blend so perfect that it was difficult for him to realize how others found contradictions in the composing elements."<sup>2</sup> Greek humanism was a subject which he studied intensively in his youth and upon which he lectured and wrote in later life. The people of ancient Greece were confronted by problems similar to those of our own age: intellectual ferment, the collapse of traditional morality, war, skepticism and materialism. In attempting to solve these problems, they created a rational system of morals and religion and a great human ideal. They thus provided us with textbooks of life which we would do well to study.

Greek philosophy, however great, did not satisfy man's spiritual hunger. The great spiritual force of Christianity complemented Hellenism and blended with it to form the basis of our culture. Livingstone was concerned about the warning influence of Christianity, and he attempted to show the need for religion in our education: to give us a proper perspective, an understanding of Western civilization (since for more than a thousand years this creed has ruled the thought and colored the conduct of the West), and a respect for human beings as *human beings*, and to assist us in the essential task of formulating a philosophy of life.

Livingstone saw the Greek and Christian ideal of democracy most nearly approximated in Denmark largely as a result of the Danish People's High School movement. He saw that these Folk High Schools have three secrets of success. First, they are open only to adults who bring to their studies an adult intelligence, a sense of the value of education and practical experience of life which make studies meaningful. Second, they are residential; thus the students become steeped in an atmosphere of education, the teacher becomes more of a personality and influence, and the pleasant buildings, gardens, pictures, music and corporate life mould the character. Third, they are spiritually oriented, "steadily insisting on a spiritual philosophy of life suited to the needs and capacities of the ordinary man."<sup>3</sup> From this movement which created the only educated democracy in the world, Livingstone drew two lessons. One was that national education should be based on adult study. The other was that national education should be made "a spiritual force, awakening and inspiring."<sup>4</sup>

### Livingstone's Philosophy

Although Livingstone emphasized the need of each individual to have a philosophy, nowhere in his writings did he make a comprehensive, explicit statement of his own philosophical position. How-

2. *The Times*, "Sir Richard Livingstone: Greek Scholar and Humanist", Dec. 28, 1960.

3. Sir Richard Livingstone. *On Education*. Cambridge University Press, 1960, p. 50.

4. *Ibid.*, p. 534.

ever, it is possible to draw together philosophical ideas from scattered statements throughout his works, as for example:

We are human beings, with a body, a mind, and also something elusive and indescribable but very real which we call a soul or spirit—three elements which combine in our personality and which interact on each other.<sup>5</sup>

The body is the most certain, tangible thing in man. It is the element in which his passions and appetites reside, irrational but capable of being obedient to reason, and having its activities and virtues which constitute man's moral life.<sup>6</sup> The intellect is "a faculty ministering to a strange need called a sense of truth,"<sup>7</sup> it is the seat of what the Greeks regarded as the highest thing in human nature.

The soul has two eyes . . . There is the critical, analytical eye which measures and assesses and protects men from illusions and delusions. But there is also the eye which enables them to contemplate, enjoy, and adore.<sup>8</sup>

For the soul is the element of the divine in man. Like his classical hero, Plato, Livingstone believed that the full and harmonious development of these three faculties will result in a happy and virtuous man. Man's fundamental purpose, then, is to make the best of human nature (i.e. to achieve the 'virtue' of man).

On the basis of the foregoing ideas, Livingstone built his system of ethics. A moral system, he believed, must be based on reason and on Nature—on man's "qualities, capacities, and uses."<sup>9</sup> The cardinal features of a great moral system are three. First, it is disinterested, "leading men to desire the good in all its forms, not for results or accidental advantages but for itself." Second, it is progressive, "developing from lower to higher conceptions of good in a vision which is progressively purified and enlarged"<sup>10</sup> as human intellect and experience grow. Third, it is free from narrowness, for the ideal to which it leads is rich and many-sided combining public and private, practical, intellectual and artistic pursuits.

Closely interwoven with Livingstone's ethical theories are his political theories, for civilization rises or falls with the quality of man's interests, tastes, desires, and capacities to which it adapts. Society must therefore collapse unless men have a quantum of moral qualities, and can be improved only by improving men. It is necessary to show them the spiritual ideals without which neither happiness nor success is genuine or permanent, to get them to accept excellence as their master in vocational, intellectual, aesthetic, civic and personal spheres of activity.

5. Sir Richard Livingstone. *The Rainbow Bridge*. Clarke-Irwin, Toronto, 1959, p. 3.

6. Sir Richard Livingstone. *Education and the Spirit of the Ages*, p. 49.

7. Sir Richard Livingstone. *The Greek Genius and its Meaning to Us*. Oxford University Press, London, 1912, p. 134.

8. Sir Richard Livingstone. *Education and the Spirit of the Age*, p. 102.

9. Sir Richard Livingstone. *Greek Ideals and Modern Life*. Oxford University Press, 1935, p. 140.

10. *Ibid.* p. 74.



As with Plato, so with Livingstone, the key to an improvement in human character and thus in politics and the world, was education. A few of his epistemological tenets deserve consideration. He distinguished between two types of knowledge which one might call knowledge of facts and knowledge with appreciation. There are thus also two ways of acquiring knowledge. One way is logical thinking and analysis. But knowledge depends upon imagination as well as reason; hence this method takes us only part of the way. The second means of arriving at truth is vision or intuition. If an education is to be adequate, it must achieve a balance between the analytic intellect which protects the mind from fantasies and disciplines it, and vision which reveals realities beyond critical analysis.

One of the reasons why much of our education fails is that young people do not want to, and cannot study fruitfully such cultural subjects as history, literature, politics, economics, and ethics because they do not have the necessary knowledge of life derived from experience. The student can acquire from such subjects a background of specific knowledge as well as a number of intellectual skills, but he cannot test, interpret, utilize or really desire them. Therefore adult education is essential.

### **Implications of Livingstone's Theories for Educational Practice**

The tripartite nature of man determines the aims of education. All men need to earn a living; therefore, education must supply them with the techniques on which our civilization is based. All men need to learn to live in a society; therefore, education must give them an understanding of the universe and of men, so that they may be intelligent citizens. And all men have personalities; therefore education should help them to develop fully as human beings so that they may live the good life. Education must do these things for *all* citizens, though the means to accomplish them will vary with different people. The first of these aims is receiving the major emphasis, but it is the latter upon which the improvement of citizens and hence of the world depends.

What curriculum would best achieve these aims? Most of our schools have experienced a crowding of the curriculum as a result of the vast increase in knowledge and try to teach everything with the result that they teach nothing well. A good education must be narrow. It must observe two principles: first, two subjects must be studied thoroughly, so that the pupil will learn what knowledge really is and what its acquisition entails; second, these subjects should give the pupil a vision of greatness. Literature and history, the record of human visions of life and the record of human actions, are best able to accomplish this end. Philosophy and religion, too,



are valuable. Philosophy "clears the mind, it forces people to think, it makes them realize the importance and value of having a fixed principle by which to test and guide and strengthen conduct and it sets them on the road to finding such a principle."<sup>11</sup> But it is religion which moves humanity and satisfies its needs. It is the supreme instrument for that education which provides ideals of conduct. Other subjects that are "charged with values" are art, music, poetry and drama. They are valuable in that they stimulate and develop the powers of imagination, feeling and appreciation and thus counteract the critical and analytical spirit of the age. Even physical education provides opportunities for character training. And Livingstone, like the Greeks whom he so admired, believed that a disciplined body should be the servant of a virtuous mind.

Having used the foregoing subjects to develop man as the best human being possible, one can group around this central core the accessories which he uses to achieve his purposes:

science, by which he increases his knowledge and control of the universe; politics and economics, by which he creates and regulates the society which will best serve the good life; languages, through which he has access not only to his fellowmen but to the collective wisdom of the world; industry and commerce, regarded not merely as a means of making money, but as Plato conceived them, as mothers and nurses that supply mankind with the necessities of life. These accessories should be seen and studied with continual regard to the central figure and not, as they too often are, treated as independent forces, each to be elaborated for its own sake without any care for the real subject of the picture.<sup>12</sup>

The teacher's task is given. He must eradicate ignorance, and in its place create reason, right conviction, and wisdom. In order to do this, the teacher must have certain qualities.

- (1) He must himself have a sense of values about which he feels intensely, even though he may never mention them, "for a teacher's outlook educates, more than anything he says."<sup>13</sup>
- (2) He must know, and be able to make his pupils see, the first-rate.
- (3) He should be able to whet the pupils' appetite for knowledge and develop their capacity to learn, rather than load their minds with "undigested lumps of knowledge."
- (4) He should have a definite view of his aim and test for its achievement by the clinical study of his students.
- (5) He should be a specialist, but not such an expert that, in devoting himself to those aspects of his subject with which the universities concern themselves, he misses the value of his subject in the training of human character.<sup>14</sup>

11. Sir Richard Livingstone. *Education and the Spirit of the Age*, p. 38.

12. Sir Richard Livingstone. *Some Tasks for Education*. Oxford University Press, London, 1946, pp. 19-20.

13. Sir Richard Livingstone. *On Education*, p. 100.

14. Sir Richard Livingstone. *The Rainbow Bridge*, p. 10.

Having obtained a view of what human beings should be, and arrived at conclusions regarding the curriculum that will produce such type, Livingstone considered the methods to be employed. For a good education depends not merely upon the subjects taught but also upon the way in which they are taught. Every method has its defects and limitations, and different methods favor different subjects and develop different virtues. The appropriateness of the different methods varies also with the ability and age of the students. Effective methods may be based on the main principles of the movement which stems from Rousseau, Rabelais and Montaigne:

The child must be active, not passive. Education is to be a process of living, not a preparation for it. The school is a place, not for giving a body of authoritative information but for developing the child's creative powers. Teaching must not outrun his experience or attempt to plant in him knowledge which has no roots in his own life. Literary study is to be reduced or even abolished.<sup>15</sup>

While Dewey is right in saying that the child should be formed by contact with the world, it cannot be merely the everyday world; rather it must be the world at its best. A sense of values must be imparted, and this is best done by showing the students the first-rate in all fields until they are satisfied with nothing less. But to teach them these ideals is not enough; to know the good is not the same thing as to do it. Hence it is important that our early education inculcates good habits by creating conditions which will automatically develop and stabilize character. There is no better place for this than the boarding school, "the best training ground for citizenship in the world."<sup>16</sup>

Analysis is an important method of education, and is receiving emphasis in our schools. It is even more important, however, to draw out habits of aesthetic appreciation, to teach people to see and feel. Associated with aesthetic appreciation, a method of great value which is not receiving sufficient attention is memorization, particularly of the Bible and of other great literature. At the time, the students will resent this method, but these treasures stored in the memory will in time be illuminated by life and in turn illuminate it. Examinations, on the other hand, "are the curse of education."<sup>17</sup> Though they are necessary to concentrate the mind and to make knowledge precise, to test industry or ability and to spur the pupils on to greater effort, they become a corrupting as well as a stimulating influence.

The most striking fact regarding educational methods is how little we really know about them.<sup>18</sup> Education, like medicine, deals with individuals no two of which can be counted upon to respond

15. *Ibid.* p. 120.

16. *Ibid.* p. 42.

17. Sir Richard Livingstone. *Education and the Spirit of the Age*, p. 106.

18. Sir Richard Livingstone. *The Rainbow Bridge*, p. 32.

in exactly the same way. We need, therefore, to be always watchful and to continually ask ourselves what is happening. In short, we must use the technique of clinical study.

Passing from the restricted field of education in its formal sense to education within the society at large, Livingstone observed that the majority of our education does not pass by that name and takes place outside the school and university. The home educates, conditions of living and work educate, newspapers, films, advertisements educate; above all, public opinion, the accepted standards of good and evil, right and wrong, educate.<sup>19</sup>

These agents of education are at least as effective as what goes on in the classroom; for people are sensitive, imitative creatures, unconsciously moulded by the atmosphere which they create. It is therefore essential that homes, living conditions, books, newspapers, magazines, movies, the theatre and television be good. Otherwise, they will defeat the purpose of the schools.

Sir Richard's standpoint on education has many attractions. Aside from the fact that its protagonist was a thinker, scholar and administrator who commanded high respect, its sanity, balance and magnanimity are impressive. We are indebted to him for his critical analysis of contemporary education, for the forceful expression of old ideas and ideals in terms of new needs, and for the vision of what life can mean to people when education enables them to get the most out of it. Few sensitive persons in the western world could read his works without feeling unlifted, enriched, and challenged to excel.

NOTE: Sir Richard's theories regarding his main field of interest, adult education including university education, are too vast and too significant to come within the scope of this short article.

---

19. Sir Richard Livingstone. *Some Tasks for Education*, p. 63.



# THE ACADEMIC INTERESTS OF FAILING COLLEGE STUDENTS IN ARTS AND SCIENCE

E. R. OETTING

*University of Alberta, Calgary*

An academic interest inventory was administered to the 1962-63 freshman class at the University of Alberta, Calgary. The test consisted of 132 pairs of statements, describing the kind of subject matter that might occur in high school or college courses. The student made a forced choice indicating his preference for one item of the pair, even though he may like or dislike both alternatives. The test was administered in freshman week after the students had registered in their faculty and programs, and before they had started classes.

At the end of the year successful and unsuccessful college students were selected on the basis of their performance in the university. Successful students were identified as those who had completed the year and were eligible to enter the second year of university work. Failing students were those who had completed the year but were dropped from the university for poor performance, required to repeat the first year, or failed to receive credit for part of their program. Students who were dropped during the year were not considered. The responses to the test items were then separately compared for the successful and failing students in each faculty.

Twenty-four items differentiated between successful and failing students in the first year of the Bachelor of Arts programs ( $X^2$  greater than .05 level). Test results were available for 94 successful and 23 failing Arts students. A summary of the types of items that differentiated between Arts students appears in Table 1. The statements in the test itself are more complete and descriptive, but the phrases in the table summarize the content of the items. An examination of Table 1 indicates considerable consistency in the type of item that was more frequently endorsed by failing students. Most of the items are technical and a number of these involve rather pragmatic practical activities. Several of the other items would involve commercial activity. Opposite these items are descriptions of course material that were selected less frequently by the failing students. Because it was a forced choice test, an item could have differentiated between the groups either because the failing students more frequently preferred one of the alternatives or because they did not like the course content described opposite it. How-

TABLE I  
ACADEMIC INTERESTS OF BACHELOR OF ARTS STUDENTS

% of failing	% of passing	Content preferred more frequently by failing than passing students	% of failing	% of passing	Content preferred less frequently by failing than passing students
44	31	Control heat loss	56	69	Cell growth
39	22	Gear ratios	61	78	Methods teaching
57	39	Landscaping	43	61	Interest students in learning
74	43	Chemical isotopes	26	57	Conjugate verbs
52	40	Taxation and finance	48	60	Folk stories and myths
57	39	Formation of jewels	43	61	Teamwork in children
55	43	Theory of numbers	45	57	Origin of words
56	43	Design canals	44	57	Paint scenes
52	27	Make metal castings	48	73	Graph statistics
48	35	Spectra of minerals	52	65	Glands of body
52	30	Friction in machinery	48	70	Geometry
70	53	Plant chemistry	30	47	Furniture fabrics
35	19	Automatic machinery	65	81	Mental illness
57	46	Analyze languages	43	54	Students personal problems
43	28	Federal Banks	57	72	Roman Empire
78	40	Probability combinations	22	60	Book report
57	30	Theory calculators	43	70	Retail store operation
52	38	Theory of gravitation	48	62	Water supply systems
57	34	Make depth soundings	43	66	Family adjustment
43	25	Social life of ants	57	75	Interview for news
52	39	Sing in choir	48	61	Use microscope
74	54	Interview unemployed	26	46	Write letters
61	43	Inspect with X-rays	39	57	Supervise playground

TABLE II  
ACADEMIC INTERESTS OF BACHELOR  
OF SCIENCE STUDENTS

% of failing	% of passing	Content preferred more frequently by failing than passing students	% of failing	% of passing	Content preferred less frequently by failing than passing students
44	30	Powers of government	56	70	Math equation
39	20	Give political speech	61	80	Foreign language
67	47	Induction coils	33	53	Logarithms
56	32	Company publicity	44	68	Music appreciation
39	24	Taxation and finance	61	76	Folk stories and myths
83	71	Fossils under microscope	17	29	Square roots
64	46	Design canals	36	54	Paint scenes
48	31	Compute averages	52	69	Babylonian civilization
75	63	Employee morale	25	37	American literature
61	40	Labor movement	39	60	Compare languages
45	21	History of architecture	55	79	Math problems
71	56	Paint pottery	29	44	Compare authors
44	27	Technique selling	56	73	Geometry
81	57	Fuse porcelain	19	43	Read foreign papers
58	44	Interview unemployed	42	56	Write foreign letters
53	35	Announce commercials	47	65	Napoleon's Law
100	66	Analyze urine	0	34	Count foreign words
67	34	Theory electron tube	33	66	Compare foreign art

ever, just as there is considerable consistency in the descriptions of the items selected more frequently, there is a great deal of consistency in the type of content that was preferred less often by failures than by successful students. A number of the items suggest that failing students are not interested in students and how they function, and other items suggest a dislike of material that might



be found in English or Social Studies courses. When faced with a choice between biological science content and a more technical science, the failing Arts students more frequently preferred the technical science.

Eighteen items differentiated between successful and failing Bachelor of Science students. Test results were available for 91 successful and 36 failing science students. The failing students in the B.Sc. patterns tended to express preferences for course content that was fairly technical and usually quite practical, rather than theoretical, as well as content that would involve activities in politics or business (Table II). Interestingly, they showed a lack of preference for several items that involved mathematics, despite election of science programs. They also expressed a preference for anything involving foreign language or literature much less frequently than did the successful students in science programs.

The tendency for failing students in both Arts and Science to indicate a preference for technical and rather practical course content may indicate one of their problems in adjusting to a college program that is oriented toward theory and underlying principles, rather than practical, applied, factual knowledge. Students in science who tend less frequently to prefer content involving mathematics may also be indicating very directly a source of academic maladjustment. The tendency of failing science students to endorse less often content that involved foreign language or English, and of failing arts students in not preferring English or Social Studies to technical content, suggests that both groups may be less broad in general interests, and may not like to read or express themselves in writing as contrasted to their more successful counterparts.

Counsellors have traditionally considered academic interests in helping students select a college major that they will like and one that will meet their future needs. The results of this study indicate that the counselor might also find it profitable to explore the student's expression of academic interest as an indication of his chances of success or failure in addition to the more usual consideration in selecting the college major. The attitudes underlying these expressions of preference are already formed before entrance to the university. It might also be profitable if we gave further consideration to the kind of influences that tend to develop highly pragmatic practical values in the matriculation students. The counselor should also be sensitive to the kind of environmental influences that lead to a rejection of certain subject areas such as Mathematics, Foreign Languages, English, Social Studies or Education, and should consider the expression of a dislike for certain types of content not only as a reflection of interest, but also as an indicator of potential problems in doing successful college work.

# DIFFERENTIAL ACHIEVEMENT IN READING AND ARITHMETIC

By Pupils of High General Ability

MARY H. NEVILLE AND BARRY P. FROST

*University of Alberta, Calgary*

## The Problem

The prediction of the achievement of any child is always of great interest to teachers. Although it is common, at the present time, to use intelligence tests for this purpose, as Mackay and Vernon (1963) say "we cannot simply state—that changes in performance following instruction are identical with some hypothesised intelligence." Ferguson (1954) following Hebb (1949) suggests that intelligence is the result of the interaction of the individual's biological characteristics and learning. He claims that the 'ability to effect positive transfer of experience to a new learning situation' is a vital factor in intellectual functioning. This 'positive transfer' is thought to be analogous to Spearman's 'g' (Spearman 1927), while specific abilities develop as the result of specific learning situations.

As Nancy Bayley remarks, "The very fact that the scores of mental growth in individual children tend to exhibit gradual shifts in relative status supports the theory that a changing organization of factors is in process. Something akin to 'g', or a high first factor loading, must appear soon after the second or third year." (Bayley 1955, p. 808).

Similarly Garrett (1946, p. 375) claims "Over the elementary school years we find a functional generality among tests at the symbol level. Later on this general factor or 'g' breaks down into the quasi-independent factors reported by many investigators. It seems likely that the 'g' factor which appears strongly at the elementary school level is, in large part, verbal and linguistic in nature. If the school child can read well, he can very probably do the rest of his school work well."

This view of 'intelligence' and specific abilities is supported by research showing the change with age of the organization of abilities which in the early grades are very little differentiated (Anastasi 1958, Tyler 1953).

The high positive correlations between the sub-tests of the Wechsler Intelligence Scale for Children (Wechsler 1949) or Thurstone's Primary Mental Abilities Test (Buros 1953, pp. 698-710) also indicate that the child's intellectual functioning is not solely a mani-



festation of separate specific factors.. More specifically, as Jastak (1949) says, "—the number factor and the verbal factor always yield scores that are positively correlated. The independence of these factor has not been demonstrated up to this point."

If this is so then we would expect children, particularly young children, to have very similar achievement in both arithmetic and reading (at least to the extent that these ability factors, referred to above, are predictive of school achievement).

There is much evidence to show that intelligence tests may be useful predictions of scholastic achievement (e.g. Littell 1960) (though one must be cautious—cf. Frost 1963), and this may be so because in the new learning situation of the school the child will function on the basis of his prior learnings which may be more highly correlated with intelligence as measured—i.e. the tests give some indication of the individual's capacity to transfer experience to the new learning situation.

In a previous study conducted by one of us (Neville 1950), it was observed that, although elementary school children with IQ's above 120 (Otis, New Zealand Adaptation) had reading ages (ACER Silent Reading Test Form C) as high or higher than their mental ages, their arithmetic ages (ACER Arithmetic Tests) were almost the same as their chronological ages. Other investigators (Barratt and Baumgarten, 1957, Cleland and Toussaint 1962, and Mitchell 1963) have reported that, while correlation coefficients of approximately 0.60 have been found between reading test scores and intelligence test scores, the correlation coefficients between arithmetic test scores and the latter were only of the order of 0.20. Frequently pupils with high IQ's were found to perform quite poorly in arithmetic while those with low IQ's performed better than predicted. As a result of these findings, Mitchell (1963) has suggested that "the achievement expected—predicted for a given level of IQ—must be established separately for each subject test on the achievement battery."

When the results of "achievers" and non-achievers" were considered separately (Barratt and Baumgarten, 1957), the correlation between WISC or Binet IQ's and reading was nearly the same for both groups, but the correlation for arithmetic and IQ depended upon the group. In the "achieving" group the correlation coefficient demonstrated an almost chance relationship whereas in the "non-achieving" group the relationship was significant. As Littell (1960) comments "this suggests strongly that other important variables are involved."

What variables may these be? One explanation may be found in Strang's statement "Gifted children often help themselves to gain



proficiency in reading," (Strang, McCullough and Traxler, 1961, p. 265). It seems probably that many intelligent children by their own efforts improve their reading ability and hence their vocabulary and language skills. Parents, too, may encourage reading by providing books and library cards, or by their own example. The same effect is much less likely to be seen in the field of arithmetic. It is true that children may speculate on problems involving numbers and perhaps improve their number concepts by their own efforts. But, while they can become more and more skillful in reading after as little as one year of basic training, in arithmetic, unless they are actually taught fundamental skills and computational methods, they will not, in general, possess the 'tools' for independent work of a high standard. At home, parents do not usually provide materials for practice in arithmetic, nor do they have the experience and knowledge to help the child increase his own knowledge systematically. Only recently have books and materials been available which are within the range of convenience of parents—e.g. the Cuisenaire System (Gattegno 1961).

In view of the above data the hypothesis which suggests itself is that children with high IQ's are retarded in arithmetic achievement relative to their "mental" and reading ages.

### Method

To overcome, as much as possible, the contamination of the intelligence test scores with the reading factor, it was decided to use a non-verbal test of intelligence. A convenient test, and one known to be highly saturated with 'g' (MacArthur 1962, MacArthur and West 1964) is the Safran Culture Reduced Intelligence Test (SCRIT) (Safran 1960). This test is normed on the Calgary population.

Consequently, this test was given to the entire Grade IV population (92 children of ages 8-0 to 11-11) of a Calgary public school in a new housing development. The socio-economic background of the children ranges from artisan up to professional with a definite weighting in the middle class.

The reading and arithmetic achievement of the children was tested by the California Achievement Battery Elementary (Grades 4, 5 and 6) Form W. Only the arithmetic and reading sections were used, the language test being eliminated. This level of the test was used to ensure that the superior pupils would not gain 'perfect scores'. The reading section of this battery is made up of two sub-tests, Vocabulary and Comprehension; the arithmetic section of Reasoning and Fundamentals.

Testing was completed within a week. Therefore all test results can be considered to be representative of performance at one given period of the children's development.

The achievement scores of the children scoring above the 85th percentile on the SCRIT were analysed as set out below and, as a follow up to the Barratt and Baumgarten study, the scores of those below the 31st percentile on the SCRIT were also analysed. Using these percentile cut-off points gave us a group of high IQ pupils (Group G) and a group of low IQ pupils (Group D). There were 20 children in group G (10 boys, 10 girls) and 19 children in group D (11 boys, 8 girls).

Using the raw data (for groups G and D separately) product moment correlation coefficients were calculated for the achievement tests and subtests; these are given in Table II.

For the 55 children aged between 9-0 and 9-11 years (29 girls, 26 boys) product moment correlation coefficients were computed for SCRIT and the Achievement tests (raw scores being used). These coefficients are given in Table III.

"High Grade IV" norms were used to convert the achievement test raw scores to percentile scores for groups G and D. These percentile norms provide a simple and clear means of comparison of each child's performance on the different tests as well as a means of comparison of individuals.

Results

The Sign Test (Binomial) was applied to the reading and arithmetic percentile distributions of group G. The computations give a probability of 0.006. Similar calculations for group D give a probability of 0.032. For both groups the reading scores are significantly higher than the arithmetic scores.

This fact is also reflected in the mean percentile scores for the achievement tests and sub-tests shown in Table I. It is obvious that the low Fundamentals Scores are the major source of the differences.

TABLE I  
MEAN PERCENTILE SCORES

	Group G	Group D
Vocabulary .....	90	80
Comprehension .....	90+	70
Total Reading .....	95	80
Reasoning .....	90	60
Fundamentals .....	50	30
Total Arithmetic .....	80	50

With the exception of the inter-correlation of the sub-tests, the correlation coefficient for the tests and sub-tests (Groups G and D) are low and insignificant as can be seen in Table II.

TABLE II  
CORRELATION COEFFICIENTS FOR READING  
AND ARITHMETIC TESTS

Tests	Group G (N=20)	Group D (N=19)
Vocabulary and Total Reading .....	0.85xx	0.90xx
Comprehension and Total Reading .....	0.84xx	0.93xx
Vocabulary and Comprehension .....	0.69xx	0.49x
Reasoning and Fundamentals .....	0.55x	0.39
Total Reading and Total Arithmetic .....	0.24	0.37
Total Reading and Fundamentals .....	0.17	0.27
Total Reading and Reasoning .....	0.32	0.36

xx—Significant at 0.01 level or better  
x—Significant at 0.05 level or better

TABLE III  
CORRELATION COEFFICIENTS FOR SCRIT  
AND ACHIEVEMENT TESTS  
(N=Girls 29, Boys 26)

	SCRIT		Reading Total	
	Girls	Boys	Girls	Boys
Arithmetic (Total) .....	0.45	0.38	0.51	0.65
SCRIT .....			0.32	0.66

TABLE IV  
FREQUENCY DISTRIBUTION OF READING AND  
ARITHMETIC PERCENTILE SCORES  
(High Grade IV Norms)  
(Groups G and D)

Percentile Score	Total		Reading Reasoning		Fundamentals		Arithmetic Total	
	G	D	G	D	G	D	G	D
90-99 .....	16	3	17	6	3	1	7	2
80-89 .....	3	7	3	3	1	0	7	3
70-79 .....	0	5	0	1	1	3	3	2
60-69 .....	0	1	0	2	3	1	3	1
50-59 .....	1	0	0	5	5	3	0	4
40-49 .....	0	0	0	2	4	2	0	2
30-39 .....	0	3	0	0	3	1	0	4
20-29 .....	0	0	0	0	0	4	0	1
10-19 .....	0	0	0	0	0	3	0	0
0- 9 .....	0	0	0	0	0	1	0	0
Σ .....	20	19	20	19	20	19	20	19



TABLE V  
RAW SCORES ON SCRIT AND ACHIEVEMENT TESTS

	Boys		Girls	
	$\bar{X}$	S.D.	$\bar{X}$	S.D.
SCRIT .....	30.50	2.78	29.28	3.60
Arithmetic (total) .....	49.00	7.16	45.62	9.15
Reading Total .....	77.31	17.54	76.10	13.19

TABLE VI  
CORRELATION COEFFICIENTS FOR SCRIT  
AND ACHIEVEMENT TESTS

SCRIT .....	Arithmetic		Reading	
	High Achievers	Low Achievers	High Achievers	Low Achievers
	0.19	0.39	0.18	0.80

Discussion

The results of the Sign test confirm our hypothesis that children with high IQ's are retarded in arithmetic achievement relative to their reading ages. The percentile scores set out in Table I confirm that they are retarded in arithmetic relative to their "mental ages" but that most of the retardation is in "fundamentals" not "reasoning."

Our results show that the same situation is also true of the low IQ group. It would seem that in this sample of children from a middle class background, irrespective of IQ level (within the "normal" range) reading attainment is considerably higher than arithmetic attainment (at least as far as fundamentals are concerned).

The correlation coefficients set out in Table III demonstrate that there is some large degree of truth in Jastak's remark quoted above. There is a significant lack of independence in reading and arithmetic scores. In fact the reading scores were as good a predictor of arithmetic attainment as the non-verbal intelligence test scores.

As far as boys are concerned our results support Barratt and Baumgarten's statement that intelligence tests predict reading better than arithmetic, but our results show the opposite to be true for girls. As can be seen from Table V, part of the reason for this sex difference is the girls' smaller variance in reading scores and larger variance in arithmetic as compared with the boys.

Splitting the 9 year old distributions (on Reading and Arithmetic separately), at the median, to give high and low achieving groups on reading and arithmetic respectively, and correlating the scores with the SCRIT, gave the results of Table VI. It can be readily seen that these figures do not confirm the findings of Barrett and Baumgarten

referred to above. Presumably the correlations discovered will vary with the intelligence test used. Another factor here is the narrow variance of the SCRIT scores at the upper levels of achievement.

It may of course, be argued that the norms for the California Achievement Battery are U.S. norms, even though they are the result of very careful nation-wide representative sampling, and that they are not applicable to this sample of children from a new housing sub-division in Calgary.

To affect our argument it would have to be the case that U.S. children were worse in reading and better in arithmetic than our children.

The "Fundamentals" test is a test of the traditional compartmentalised arithmetic and from grade III onwards our children have followed a new arithmetic course largely influenced by set theory (Hartung, Van Engen and Knowles, 1955). Tests have shown, however, that children following this new course are, if anything, better in computational skills than those following the old course. (Lindstedt, 1963). If a Grade IV child is able to add, subtract, multiply and divide correctly, using whole numbers only, his score will fall on the 95th percentile of the California Test norms.

It seems possible, as Harris (1961) has pointed out, that mediocre performance of intelligent children in some aspects of arithmetic may be due to a type of teaching which aims at uniformity of achievement in a class. In the teaching of reading there is some recognition in the schools, of the tremendous range in level of performance and rate of change (Olson, 1957; Frost, 1963). The same teachers teach arithmetic but do not seem to allow for variance in performance in this subject as they do for reading. Yet it is here even more than in reading that special teaching and materials should be made available to the able learner. Although a special program would be of great help to him in reading, the able learner may partly work out his own development by the very nature of our "school culture" which is biased verbally and provides for the middle class child many opportunities for sophistication of reading. He does not have the same natural opportunities to develop mathematical ability and yet relative innumeracy is, in the 20th century technological culture, as important as relative illiteracy.

There has been a good deal of discussion about the question of a 'ceiling' to the individual's ability to learn mathematics (Bateman 1960), and it is commonplace to hear of otherwise intellectually competent children, who cannot 'do' mathematics. It is difficult to see why a child who functions at a high level in other cognitive activities should necessarily be impotent in the field of numbers.



New approaches to arithmetic whether of the "structural" type (Stern, 1949, Cuisenaire and Gattegno, 1954, Gattegno 1960, Dienes 1960, Williams 1961, 1962) or of the "pattern of thinking" kind (Hartung, Van Engen and Knowles 1955, Lindstedt 1963), with their stress on the understanding of arithmetical processes and their break with compartmentalisation in favour of an integrated approach, should help greatly to improve the performance of the "innumerate" child. However, the correlation which exists between the subtests of "Reasoning" and "Fundamentals" should remind us of how dependent the child is on the "tools" of computation, and warn us not to overlook the systematic development of, and practice in, these skills (cf. Bruner, 1960).

This investigation is confined in its generality by a) the size of the samples and b) the socio-economic background of the children concerned. Similar results might not be found in children from a working class background where the environment may be much less favorable to the development of reading ability.

It would be very useful to replicate this study using a) larger N's, b)) a wider range of socio-economic backgrounds, c) a wider range of modes of teaching and d) tests normed on the local population.

#### REFERENCES

- Anastasi, Anne. *Differential Psychology*, 3rd ed. New York, Macmillan 1958.
- Barratt, E. S. and Baumgarten, D. L. "The Relationship of the WISC and Stanford Binet to School Achievement." *Journal of Consulting Psychology*, 1957, 21, 144.
- Bateman, E. H. "Mathematics and the Structural Engineer." *The Structural Engineer* 1960, 38, 123-128.
- Bayley, Nancy. "On the Growth of Intelligence." *The American Psychologist* 1955, 10, 805-818.
- Bruner, J. S. *The Process of Education*. Cambridge, Mass., Harvard U.P. 1960.
- Buros, O. K. *The Fourth Mental Measurements Yearbook*. Highland Park, N.J. Gryphon Press 1953.
- Cleland, D. L. and Toussaint, I. H. "The Interrelationships of Reading, Listening, Arithmetic Computation and Intelligence." *The Reading Teacher* 1962, 15, 228-232.
- Cuisenaire, G. and Gattegno, C. *Numbers in Color*. London, Heinemann 1954.
- Dienes, Z. P. *Building up Mathematics*. London, Hutchinson 1960.
- Ferguson, G. M. "On Learning and Human Ability." *Canadian Journal of Psychology*, 1954, 8, 95-112.
- Frost, B. P. "The Role of Intelligence 'C' in the Selection of Children for Remedial Teaching of Reading." *Alberta Journal of Educational Research*, 1963, 9, 73-78.
- Garrett, H. "A Developmental Theory of Intelligence." *The American Psychologist*, 1964, 1, 372-378.
- Gattegno, C. *A Teacher's Introduction to the Cuisenaire-Gattegno Method of Teaching Arithmetic*. Reading, Gattegno-Pollock Co. 1960.
- Gattegno, C. *Now Johnny Can Do Arithmetic*. Vancouver. Cuisenaire System of Canada Ltd., 1961.



- Harris, A.J. *How to Increase Reading Ability*. New York, Longman, Green, 1961.
- Hartung, M. L., Van Engen, H., and Knowles Lois. *Seeing Through Arithmetic*. Chicago Scott, Foresman 1955.
- Hebb D. O. *The Organization of Behavior*. New York, Wiley, 1949.
- Jastak, J. "Problems of Psychometric Scatter Analysis." *Psychological Bulletin*, 1949, 46, 177-197.
- Lindstedt, S. A. "Changes in Patterns of Thinking Produced by a Specific Problem Solving Approach in Elementary Arithmetic." *Alberta Journal of Educational Research*, 1963, 9, 65-72.
- Littell, W. M. "The Wechsler Intelligence Scale for Children: Review of a Decade of Research." *Psychological Bulletin*, 1960, 57, 133-156.
- MacArthur, R. S. *The Construct Validity of a Number of Measures of Intellectual Potential for an Alberta Metis Sample*. (Paper presented to the Annual Meeting of the Canadian Psychological Association, Hamilton, June 1962).
- MacArthur, R. S. and West, L. W. "An Evaluation of Selected Intelligence Tests for Two Samples of Metis and Indian Children." *Alberta Journal of Educational Research* 1964, 10, 17-27.
- Mackay, G. W. S. and Vernon, P.E. "The Measurement of Learning Ability." *British Journal of Educational Psychology*, 1963, 33, 177-187.
- Mitchell, B. S. "A Comparison of the Achievement-Intelligence Relationship for Pupils, with that of School Systems." *Journal of Educational Research*, 1963, 57, 172-180.
- Neville, Mary H. *Intelligence and Achievement Tests for Standard Three Children*. Unpublished manuscript. Victoria College (University of New Zealand) 1950.
- Olson, W. C. *Psychological Foundations of the Curriculum*. Paris. UNESCO, 1957.
- Safran, C. *The Safran Culture-Reduced Intelligence Test*. Calgary, Calgary Public School Board, 1960.
- Spearman, C. *The Abilities of Man*. London, Macmillan, 1927.
- Stern, Catherine. *Children Discover Arithmetic*. New York, Harper, 1949.
- Strang, Ruth, McCullough, C. M. and Traxler, A. E. *The Improvement of Reading*. New York, McGraw-Hill, 1961.
- Tyler, Leona E. "Changes in Children's Scores on Primary Mental Ability Tests over a Three Year Period." *The American Psychologist*, 1953, 8, 448-449.
- Wechsler, D. *Wechsler Intelligence Scale for Children*. New York. Psychological Corporation 1949.
- Williams, J. D. "Teaching Arithmetic by Concrete Analogy—I: Minning Devices." *Educational Research*, 1961, 3, 112-125, 195-213.
- Williams, J. D. "Teaching Arithmetic by Concrete Analogy—II: Structural Systems." *Educational Research* 1961, 4, 163-192.

# AN INVESTIGATION TO IDENTIFY CREATIVITY IN TEACHING

BY DAVID J. CHABASSOL AND GORDON G. MANSON  
*University of Victoria*

## Statement of the Problem

A survey of the literature reveals an almost overwhelming number of studies which have endeavored, in one way or another, to set up research designs which would predict success in teaching. An exhaustive article by Anderson and Hunka (1963) provides a clear picture of the many attempts in this area. The majority of such studies use, as criteria for success in teaching, grades assigned by principals, supervising teachers, other administrators, and faculty of education personnel. Such grades, in turn, are influenced by a multitude of factors, which include ability to maintain discipline, the establishment of effective classroom routines, good questioning techniques, ability to present material logically and clearly, and many more. Conspicuous by its absence in this list of conventional criteria is any specific reference to creativity in teaching, that is, teaching characterized by relatively greater productivity of ideas, perception of relationships, flexibility in behavior, recognition of implications, and originality and imagination. Perhaps the single outstanding feature characterizing this variety of response is that of divergence from common, conventional, or accepted modes of behavior, as elaborated upon in Guilford's (1959) article.

Can a teacher be said to be truly successful if he is not creative? Do we expect to find in the performance of a successful teacher just a repetition of the techniques and methodology which are specifically taught him in a teacher training institution? One would hope that something more than this is necessary; that certain of those elements of creativity, as defined above, would be a *sine qua non* for success in teaching.

The present study was undertaken with the hope of providing a procedure by which creativity in practice teaching could be identified. More specifically, it was the aim of the writers to devise a practice-teaching rating scale which correlated sufficiently with the tests provided by Getzels and Jackson (1962) as to have substantial predictive value in the identification of creative teachers. It was their contention that, while the Getzels and Jackson tests were not perfected for this purpose, they had been standardized on a norm group which were not too dissimilar in age and I.Q. from the group of students used in the present study. Also, the tests of Getzels and Jackson were judged to be the best tests of creativity available for this particular group.

### **Method: Explanation of the Study**

The subjects were 100 females enrolled in the second year of the teacher training program (elementary) at the University of Victoria. All had attended the university the previous year, and had taken a five course, or fifteen unit offering.

The second year of the teacher training program at the University of Victoria is the professional or basic teacher training year. Once again, fifteen units are required, and all courses but one, namely English 200, are offered by the College of Education. These second year courses are, for the most part, methodology courses in Education.

The subjects were administered a battery of creativity tests, comprised of Word Association, Uses, Fables, and Make-Up Problems tests, as perfected by Getzels and Jackson. Two different testing sessions, of about 45 minutes each, were allotted for the administration of these tests. The content of the Problems test had been reduced from four to two problems in order to accommodate it to the time available for testing. Data available for each student included subtest score results on the Getzels and Jackson tests mentioned above and a Total Creativity Score, made up from the total of the above four subtests.

In an effort to determine the extent to which the subjects were actually giving evidence of creativity in their practice teaching, a Creativity in Practice Teaching scale was devised. The content of this scale was, in large part, based on suggestions made by instructors in the various methodology courses who, with guidance from the authors of this article, undertook to suggest how creativity in the teaching of elementary language, science, social studies, and arithmetic might be noted. It was pointed out to these instructors that creativity, for the purpose of this study, was assumed to be characterized by divergence from conventional modes of teaching and all that such divergence implies, as explained above.

Within this frame of reference, the various methods instructors then provided a variety of criteria, in the form of specific teaching practices, which they believed to be indicative of creativity in teaching social studies, science, arithmetic, and language. The resulting Creativity in Practice Teaching scale contained sixteen specific criteria which were to be graded on the basis of one point if the characteristic was "present" in the teaching of the student, and two points if the characteristic was "present in marked degree." These assessments were made over a period of four months, during which time the students were involved in five weeks of actual teaching. Any student whose teaching had been evaluated less than three times was eventually dropped from the investigation. The modal



number of observations was four, with a range of from three to seven. On the basis of the number of investigations, a mean Creativity in Practice Teaching score was determined for each student.

Additional data on each student included the following: SCAT percentile (total score); academic average on the final (April) examinations; and practice teaching grade. The practice teaching grade represents a score based on the assessment of the teaching ability of the student over the period of one academic year. Such a score is derived from a carefully planned assessment form used by both Faculty of Education members and critic teachers. Approximately eighteen or twenty observations enter into the derivation of the final score.

### *Analysis of Results*

Correlations were obtained for a variety of variables, as shown in Table I.

1. The intercorrelations among the Getzels and Jackson creativity measures show only limited agreement with the corresponding intercorrelations obtained by Getzels and Jackson in their study (1962). One is not surprised to find that a greater discrepancy occurs for the correlations involving the Make-Up Problems test which the writers reduced by one half in order to accommodate the demands of time. Omitting these latter correlations, the other coefficients, while positive, failed to reach the level of confidence attained in the Betzels and Jackson study.

This discrepancy is explained in part on the grounds that the Getzels and Jackson study figures were obtained from a population sample with a very high mean I.Q.

2. The nil correlations between the creative teaching average (item 6) and both subtest and total creativity scores (items 1, 2, 3, 4, and 5) on the Getzels and Jackson instrument are not readily explained. In any event, the purpose to which the writers applied the Getzels and Jackson tests was not fulfilled. Possibly the application of a "molecular" approach to the assessment of creativity offers only limited promise at best in the measurement of creativity in a more complicated, involved, or "molar" activity as practice teaching. The possibility also exists that the creativity in practice teaching scale devised by the authors is not a valid instrument.
3. The correlation between SCAT scores (item 7) and both subtest and total creativity scores of the Getzels and Jackson tests (items 1, 2, 3, 4, and 5) are significant at the .01 level. This is in keeping with the statement of the test makers that while creativity and intelligence are not the same thing they do tend to be positively correlated.

TABLE I  
INTERCORRELATIONS AMONG CREATIVITY  
AND OTHER VARIABLES

Variable Number	Variable	2	3	4	5	6	7	8	9
1	Word Association Test .....	.240*	.168	.314**	.604**	.017	.336**	.130	.012
2	Uses Test .....		.165	— .015	.513**	.007	.291**	.074	.058
3	Fables Test .....			.031	.308**	— .063	.259**	.253**	— .014
4	Make-up Problems Test .....				.776**	.018	.281**	.257**	.001
5	Total Creativity Scores .....					— .008	.431**	.257**	.025
6	Creative Teaching Average .....						— .057	.037	.629**
7	SCAT Total Score .....							.425**	— .015
8	April Exam Average .....								.449**
9	Practice Teaching Average .....								

4. The correlations between April examination averages (item 8) and both the subtest and total creativity scores (items 1, 2, 3, 4, and 5), while generally lower than the coefficients reported by Getzels and Jackson, tend to support their findings.
5. The correlations between the practice teaching average (item 9) and both the subtest and total creativity scores (items 1, 2, 3, 4, and 5) are nil. This comes as no surprise after considering the finding dealt with in paragraph 2, above. One might suspect that the same or a similar reason might be advanced to explain the lack of correlation.
6. The correlation between creativity in practice teaching (item 6) and practice teaching averages (item 9) is highly significant. Possibly the most obvious explanation is that the form used in assessing practice teaching at the University of Victoria places emphasis on creativity in some or all of its aspects, even though the term "creativity" is not specifically included in the criteria listed for assessment. Another explanation resides in the possibility that the instructors, being required to rate certain lessons both for creativity and for excellence in practice teaching in general, tended to permit their rating on the latter to influence their assessment of the former. However their influence was probably minimized by the fact that the overall practice teaching score was derived from as many as 20 ratings obtained over the course of the academic year, whereas, as noted above, the creativity in practice teaching score was based on approximately four ratings only.
7. The correlation between creativity in practice teaching (item 6) and SCAT scores (item 7) is very low. This low correlation and the insignificant correlation noted between SCAT scores (item 7) and practice teaching average (item 9) may be understood if one considers the contrast between the kind of performance evaluation represented by the SCAT and that undertaken in the assessment of practice teaching. The former represents a pencil and paper approach, whereas the latter is more susceptible to the subjective limitations of observational procedures. A more pertinent explanation, perhaps, rests in the discrepancy between the items evaluated by the two measures. The former presumes to evaluate intellectual functions in the broad areas of verbal and mathematical manipulation, whereas both kinds of practice teaching assessment relate to the examination of such things as adequacy of lesson preparation and presentation, assignment and supervision of seatwork, appearance, appropriate vocabulary, motivational techniques, and maintenance of discipline, to mention a few.



8. The low correlation between creativity in practice teaching (item 6) and April examination average (item 8) is not surprising inasmuch as the latter, concerned largely with educational methodology, would tend to stress convergent thinking abilities.
9. On the other hand, the significant correlation between April examination average (item 8) and practice teaching average (item 9) is not unexpected. Again, since the majority of courses taken during this year feature methodology, one might expect a high correlation between grades on these courses and marks made in practice teaching which, presumably, reflects the extent to which such methodology is being put to use in the classroom. Moreover, students are encouraged to apply in the teaching situation what they are taught in their methodology courses so that a correlation other than the nature of the one found here would be most surprising.
10. The correlation between SCAT score (item 7) and April examination average (item 8) is in the expected direction since correlations between academic potential and academic achievement tend to run between .40 and .60.

### Conclusions

What conclusions may be drawn from these data? It should be pointed out that several of the findings might have been predicted. For instance, the subtest and total test scores on the Getzels and Jackson tests correlate significantly, as was reported by their authors. Likewise, the agreement between achievement in examinations and academic ability is about what one would have expected.

However, the major purpose of this study was not to reinforce past findings but to produce a scale which would identify creativity in student teachers. If the assumption used in this research, namely that the Getzels and Jackson tests do in fact measure creativity for our population, was a correct one, then the aim of the study was not realized. Several reasons might be advanced for the complete absence of a correlation between total creativity scores as provided by the Getzels and Jackson tests and creative teaching averages provided by the instrument prepared by the authors.

1.) The possibility must be entertained that the scale devised by the authors for the measurement of creativity in teaching is not a valid instrument. Such invalidity might originate in weaknesses in construction or in usage. The former is regarded by the authors as a less likely possibility. The scale was carefully prepared, and the selection of items was felt to be consistent with contemporary views as to what constitutes creative behavior. However, with reference to usage, some possible weaknesses might be cited. It is probable

that the scale was not used with the same degree of discrimination by every rater. Furthermore, the fact that the observer was required to rate the student teacher for general teaching performance as well as for creativity in teaching at the same time might well have resulted in a halo effect. That is, the good teacher automatically becomes the creative teacher even though there are many elements influencing the former assessment which should not necessarily influence the latter. This would explain the high correlation (.629) found between these two scores.

2.) A further explanation for the lack of agreement between Getzels and Jackson test scores and creativity in practice teaching scores lies in the possibility that both of these instruments were in fact measuring creativity but in substantially different ways. As noted above, the Getzels and Jackson tests assess creativity with reference to extremely specific areas of ability, i.e., elaboration, flexibility, variety of production of ideas—in short—divergence, whereas the scale employed by the authors uses an approach which is, basically, much more general. That is, whereas the former is essentially molecular in its attack on the problem, the latter is basically molar.

3.) A final explanation for the lack of agreement between these two purported measures of creativity must be considered. The possibility exists that the scale devised by the authors does in fact measure creativity in practice teaching while the tests supplied by Getzels and Jackson do not. It should be pointed out that the Getzels and Jackson instruments were not specifically devised for the purpose to which they were put in the present research. As was noted above, the authors selected these tests, not because of their proven ability in the area of practice teaching, but because of their latent promise.

It is perhaps pertinent to note that the Getzels and Jackson tests have never been standardized on persons who have achieved success in creative endeavors. Getzels and Jackson apparently would have us believe that the individual who does well on these tests is necessarily creative because, by their definition, these are tests of creativity. A remark by Burt (1962, p. 295) is of interest at this point. In his review of the text, *Creativity and Intelligence*, in which Getzels and Jackson introduce the tests used in the present study, Burt notes, “. . . Getzels and . . . Jackson . . . apparently expect us to accept their tests as both reliable and valid without any direct or objective evidence . . .” If, in fact, these tests are not valid, the validity of the assessment of creativity in practice teaching scale devised by the authors is strongly suggested.

### Suggestions for Further Research

What suggestions could be made for the guidance of anyone who might wish to investigate in this area in future?

The writers recommend that the ratings of creativity in practice teaching and of general excellence in practice teaching not be done at the same time. As noted earlier, the grade given on the latter might influence the assessment of excellence in practice teaching in most teacher training institutions. Students feel, correctly or otherwise, that they must conform to what they consider to be expected teaching behavior. Consequently, any creative propensities might, in effect, be suppressed in the desire to please the faculty observer.

The above observation leads to the further recommendation that creativity in teaching be assessed when teachers have finished their professional year of teacher training and are operating in their own classrooms. Presumably, at such a time they would be relatively free to do as they wish, and the need to conform to someone else's conception of good teaching practice would be absent. In short, they could be as creative as they wished without fear of censure.

It is further suggested that more time and direction be given to those people carrying out the assessments of creativity in the classroom. The typical college of education faculty member, excellent though he may be in his own field, is not necessarily equipped to assess creativity in others without considerable prior instruction as to what is involved.

### REFERENCES

1. Anderson, C. C. and Hunka, S. M. "Teacher evaluation: Some problems and a proposal." *Harvard Education Review*, 1963, Vol. 33, pp. 74-95.
2. Burt, Sir Cyril. "Critical notice: The psychology of creative ability." *British Journal of Educational Psychology*, 1962, Vol. XXXII, Part 3, pp. 292-298.
3. Getzels, J. W. and Jackson, P. W. *Creativity and Intelligence*. New York, Wiley, 1962.
4. Guilford, J. P. "Three faces of intellect." *American Psychologist*, 1959, Vol. 14, pp. 469-479.



# SOME EFFECTS OF MENTAL HEALTH INSTRUCTION ON CHILDREN'S MANIFEST ANXIETY SCALE SCORES<sup>1</sup>

METRO GULUTSAN

*University of Alberta, Edmonton*

Mental hygienists are turning to the schools for help in their attempt to reduce the social and psychiatric morbidity which, exclusive of minor maladjustments, is expected to afflict one out of eight pupils now in school. Since it is impossible to alter significantly this proportion by working with individual children in clinics, there is need for preventive action with groups of children in school when children are in their formative years. It has been proposed by the Committee on Preventive Psychiatry (1951) that direct instruction in mental health might provide such preventive action.

To be effective, mental health instruction should go beyond relaying information; it should try to prevent emotional disturbances. It should help boys and girls to cope with the destructive effects of emotional conflicts by increasing their skills in managing these conflicts. To teach these skills—if it is possible to teach them—the teacher needs an established, tested program.

The specification of an instructional program in mental health is complicated by the lack of clarity as to what is 'mental health.' Marie Jahoda writes, in the introductory paragraph of *Current Concepts in Mental Health*, "There is hardly a term in current psychological thought as vague, elusive, and ambiguous as the term 'mental health'" (Jahoda, 1958). Yet, there is widespread agreement that mental health is worth promoting in the classroom. Evidence that educators think so may be found in the more recent courses of study, texts, and instructional side concerned not only with guidance and health but also with social studies, language arts and expressive arts.

Instruction in the experimental projects reported by Rosenthal (1952), Ojemann (1953), and Seeley (1954) includes the following:

- (a) permissive and unstructured discussion of problems suggested by the pupils;
- (b) role-playing or sociodrama;
- (c) catharsis, or the expression of feelings that are troublesome;
- (d) materials which emphasize the dynamic, as contrasted with the surface or superficial, approach to behavior;

---

1. This article is based on a Ph.D. dissertation submitted in 1962 to the Graduate Division of the University of California, Berkeley. The author wishes to express his appreciation for the guidance received from his advisor, Dr. F. T. Tyler. Acknowledgment is made of the valuable cooperation received from (a) the teachers of the Moose Jaw, Canada, schools who participated in the study and (b) the superintendents of schools, Mr. D. C. MacFayden and Mr. A. E. Peacock.

- (e) stories which stimulate discussion of human relations and encourage pupils to relate similar experiences of their own; and
- (f) stories with characters which serve as models for pupil identification.

Classroom instruction is under the direction of the teacher, and the approach to health instruction is thus closely related to the teacher's attitudes toward human behavior. A teacher who is sensitive to and has interest in people may add to the effects of the special instruction by increasing the pupils' skills in managing emotional conflicts. On the other hand, a less sensitive teacher may unknowingly evoke emotional conflicts in pupils without demonstrating the skills for managing such conflicts. Further, teachers in demonstration programs have probably been personally attracted to mental health projects, or have been specially selected and trained, the initial attitudes of these teachers may have contributed in significant ways to the outcomes. In the present study, the writer examines the effects of teachers with differing conceptions of human behavior but working with the same program.

A central problem in mental health is the management of anxiety. It is essential that the mentally healthy person must be able to tolerate, experience, and even anticipate anxiety (Jahoda, 1958); and for this reason the mere absence of anxiety is no longer regarded as a satisfactory index of mental health. However, because the manner a person uses for managing anxiety is difficult to determine, levels of anxiety continue to be used as rough indicators of mental health. Kitano (1960) is among the recent investigators to use levels of manifest anxiety as indices of mental health among school children. Also high levels of manifest anxiety have been found to be characteristic of:

- (a) children treated at mental health clinics, especially those diagnosed as extremely disturbed (Rynerson, 1957),
- (b) children displaying difficulties in learning when highly motivated (Sarason and associates, 1960), and
- (c) children in adjustment classes or on waiting lists for adjustment classes (Kitano, 1958).

It seems reasonable, therefore, to the present investigator that teachers with differing attitudes toward human relations may affect differently the manifest anxiety level of pupils; and they may affect differently the outcomes of the instruction they undertake.

### Procedure

Data were collected in order to obtain evidence for or against the hypothesis that a mental health instructional program will influence the manifest anxiety of children and that this influence will also



depend on the teacher's attitude to human relations. Test and re-test data on the children's manifest anxiety scale (CMAS) were obtained for grades four, five and six in both the experimental and control classes. The scale was administered by the classroom teachers on the Monday morning beginning, and the Friday morning ending, the ten-week period. Data for the teachers in these classrooms were obtained on the SM scale, by the writer.

### *Subjects and Design*

Forty-two teachers and 840 pupils from fourth, fifth and sixth grades in Moose Jaw, Canada, were the subjects of the investigation. Originally, the number of pupils was slightly larger, but some were omitted in order to provide for equal cell size and computational convenience in the  $3 \times 2 \times 2$  analysis of variance design. Each cell contained 70 cases drawn from seven classrooms. Cases were excluded by the use of random numbers. Manifest anxiety scores were obtained for all pupils at the beginning and at the end of a ten-week experimental period.

The teachers were divided into three groups according to their scores on the *Social Maturity Scale* administered about three weeks before the beginning of the experiment. Half the number of teachers in each group were randomly assigned to the experimental group and were requested to undertake the special instruction in their classrooms. All teachers so assigned agreed to carry out the special program. The remaining teachers and their pupils formed the control group in the study.

### *Meetings with Teachers*

The writer met with the teachers shortly after the opening of the fall term. Arrangements for teachers to leave their schools for the afternoon meeting were made by the two superintendents of the Moose Jaw area schools. At this meeting, the writer described the nature of the study, sought the teachers' co-operation, instructed them on the administration of the *Children's Manifest Anxiety Scale*, and administered the SM (Social Maturity) Scale to the teachers.

The twenty-one teachers who were asked to teach the special unit in mental health were given a list of readings (Jennings, 1950; Ojemann, 1953; Ross, 1957; Saskatchewan Department of Education, 1956, p. 17; Seeley, 1954; and Shacter and Bauer, 1948, pp. 17-24), and attended three ninety-minute meetings during regular school hours. The first meeting, held on the first day of the experimental period, was devoted to the introduction of readings, the sequence of instructional periods, and the procedures for role-playing and for



conducting the unstructured and permissive discussions of problems in human relations suggested by the children. The second meeting, held at the end of the second week, was devoted to the film shyness (National Film Board of Canada, 1952) and, in particular, the episode on human relations discussions. The third meeting, held at the end of the fifth week, was devoted to a discussion of the teachers' own experiences with the program.

After they had had the opportunity to try the special procedures they observed in their own classrooms the writer's demonstration of the relatively permissive, unstructured, and group-directed discussion procedure. Further, the writer conferred with each teacher on two occasions; once immediately after the classroom demonstration; and again during the sixth week of the ten-week experimental period. The writer interviewed each teacher regarding progress with the procedures, but it was not feasible to keep records of these conferences.

### *Instruction Undertaken by Teachers*

Each teacher received a guide<sup>1</sup> consisting of general information, a schedule, and a set of instruction forms with spaces for brief reports. He was asked to allot to the project the two half-hour periods each week that would normally be devoted to health instruction. As the instructional material was related to the content of the Saskatchewan curriculum guide for health, school time for it was readily available.

The instruction forms, one for each period, contained a brief description of the type of instruction to be undertaken during the period and sufficient space for brief reports and comments from the teachers. The completed forms were gathered at meetings and during class visits. Thus, there was a check on the progress of the lessons in the schools.

### **Analysis of Data**

#### *Analysis of Children's Manifest Anxiety Scale Scores*

In Table I means of CMAS scores at the time of initial testing are classified according to sex of pupils and teachers' social maturity ratings. The marked differences between the scores of girls and boys reported for their standardization sample were also found to characterize the sample in the present study.

The discovery that children with teachers scoring low on the SM scale tended to score higher on manifest anxiety is of more than passing interest. The differences between the scores of children with different types of teachers are significant at the .01 level (See Table 2). The nature of the relationship between teachers' attitudes

---

1. Available from author.

and children’s manifest anxiety is not, however, the central issue in the present study. It is not known whether teachers who are low in social maturity tend to be assigned to classrooms where children are more anxious or whether teachers with differing attitudes affect differently the level of anxiety in children. The teachers’ knowledge of pupils prior to the experiment was not surveyed and, therefore, no valid inferences are possible.

TABLE I  
MEANS OF CMAS SCORES AT BEGINNING OF EXPERIMENT

Group	Teachers’ Placement on the SM Scale				
	Upper Third M	Middle Third M	Lower Third M	Whole Group M	Whole Group S.D.
Boys .....	14.4	13.8	16.9	15.0	7.2
Girls .....	17.4	17.9	18.4	17.9	8.1
Both Sexes .....	15.7	15.9	17.7	16.5	7.8

TABLE II  
ANALYSIS OF VARIANCE OF CHILDREN’S MANIFEST ANXIETY SCORES OBTAINED BY 840 GRADES 4, 5, AND 6 CHILDREN AT START OF MENTAL HEALTH INSTRUCTION

Sources of variation	Degrees of freedom	Mean Square	F	F	F
				.05	.01
Instructional grouping .....	1	175	3.02	3.84	6.66
Type of teachers .....	2	329	5.68	3.00	4.62
Sex of pupils .....	1	1836	31.69	3.85	6.66
Within cells .....	835	57.94			

The means for the changes in CMAS scores for each of the cells used in the analysis of variance are given in Table 3.

A report of the analysis of variance of difference scores on *Children’s Manifest Anxiety Scale* appears in Table 4. It should be noted that the analysis is based on the scores of the individual pupils and not on the means of groups of children.

The only statistically significant result in Table 4 is the one associated with differences in instruction. In fact, the F ratio was large enough to be considered significant at beyond the .005 level of confidence. No statistically significant ratios were obtained for the types of teachers or the sex of pupils. It may be inferred, therefore, that the type of instruction that formed the mental health

TABLE III

MEANS FOR CHANGES IN THE CMAS SCORES OF BOYS AND GIRLS UNDER THREE KINDS OF TEACHERS AND EITHER RECEIVING OR NOT RECEIVING SPECIAL INSTRUCTION IN MENTAL HEALTH (N=70 PUPILS PER CELL)

Type of Treatment	Sex of Pupils	Teachers' Placement on the SM Scale			
		Upper Third	Middle Third	Lower Third	Whole Group
Special Instruction	Boys	1.54	.10	.10	.63
	Girls	— .11	.89	— .13	.21
B & G		.71	.49	— .01	.40
No Special Instruction	Boys	— 1.27	— .56	— .99	— .94
	Girls	— 1.63	— .11	.37	— .46
B & G		— 1.45	— .36	— .31	— .70
Whole Group		— .37	.08	— .16	— .15

TABLE IV

ANALYSIS OF VARIANCE OF DIFFERENCE SCORES IN CHILDREN'S MANIFEST ANXIETY SCORES

Source of Variation	Degrees of Freedom	Mean Square	F	F .05
Instruction (I)	1	280.01	8.16	3.85
Teachers (T)	2	20.34	.59	3.00
Pupils' Sex (S)	1	5.73	.17	3.85
I x T	2	44.68	1.30	3.00
I x S	1	26.56	.77	3.85
T x S	2	62.81	1.83	3.00
I x T x S	2	21.66	.63	3.00
Within Cells	828	34.30		

program in the present experiment, and developed on the basis of previous work in the area of teaching specifically for mental health in schools, did influence the amount and direction of change on a measure of manifest anxiety. It may be observed from Table 3, that *manifest anxiety scores tended to increase in those classrooms in which a program of special instruction was carried out, and tended to decrease in those classrooms that formed the control group and had no special instruction.* The changes of .40 and -.70 for the experimental and control groups shown in Table 3, indicate an increase of 2.5% and a decrease of 4.1% respectively. There were no significant differences in change scores for the sexes or for the three kinds of teachers.



The findings are at variance with those the writer initially anticipated. As expected, the changes were greatest for those pupils whose teachers were high on the SM scale, but contrary to expectations, the changes for the group receiving no special instruction were greater than those for the experimental group. The changes were smallest for those children whose teachers were low on the SM scale and who received the special instruction. In short, the changes were not in a mental health direction if a decrease or lowering of manifest anxiety is taken as a criterion of mental health. It has been observed by Rogers (1957, p. 95) that anxiety often increases during psychotherapy. He states, "Anxiety is often seen in therapy as the individual approaches awareness of some element of his experience which is in sharp contradiction to his self concept." It appears that increasing one's sensitivity to, and awareness of, human relations may contribute to an increase in manifest anxiety as part of the process of personality integration.

The decrease in manifest anxiety noted for the control classes was larger in absolute terms than the increase noted in the experimental classes. The decrease may reflect increased stability in children arising from increasing familiarity with the school, classroom and peer-group environment. As the children become better acquainted with the environment, and find satisfying relationships in it, their anxiety may go down. Perhaps the instructional program made the children more willing to admit to an anxiety, a possible alternative explanation in view of the fact that anxiety was identified by a self-report inventory. It would be interesting to know whether the initial and final scores in other studies using the CMAS over a period of time show changes in a uniform direction. Manifest anxiety as measured by these tests may not be the same as anxiety regarded as unhealthy. Also, a child who admits to more manifest anxiety may, in some instances, be striving toward improving his mental well-being.

The mental health instructional program did have a significant effect on the extent and direction of changes in CMAS scores. Changes in CMAS scores for children with teachers differing in attitudes were not significant. Initial CMAS scores, however, were significantly higher for those children whose teachers scored in the lower third on the *Social Maturity Scale*. It should not be inferred, however, that teachers' attitudes are unimportant variables in a mental health instructional program. Some interesting and significant differences were detected in the way high and low scoring teachers approached, and reacted to, the mental health instruction (Gulutsan, 1962).

### Conclusions

An analysis of variance design was used to test the effect of ten weeks of mental health instruction on the size of changes in Children's Manifest Anxiety Scale (CMAS) scores of intermediate grade children taught by teachers who differed on a measure of social maturity. The instruction did have an effect, significant at the .005 level, on CMAS score changes; however, the effects were not different for the different teachers. Manifest anxiety increased in children receiving the instruction and decreased in the control group. Children had higher initial CMAS scores, significant at the .01 level, in classrooms whose teachers were low in social maturity.

### REFERENCES

- Castaneda, A., McCandless, B. R. and Palermo, D. S. "The children's form of the manifest anxiety scale." *Child Development*, 1956, 27, 317-326.
- Group for the Advancement of Psychiatry, Committee on Preventive Psychiatry. *Promotion of mental health in the primary and secondary schools: An evaluation of four projects, report no. 18*, Topeka, Kansas: Author, 1951.
- Gulutsan, M. *Some effects of mental health instruction on Children's Manifest Anxiety Scale scores*. Unpublished doctoral dissertation, University of California, Berkeley, 1962.
- Jahoda, M. *Current concepts of positive mental health*. New York: Basic Books, 1959.
- Jennings, H. H. "Sociodrama as educative process." In Association for Supervision and Curriculum Development. *1951 Yearbook: Fostering mental health in our schools*, 1950, 20-285.
- Kitano, H. L. *Anxiety and rigidity in adjustment class children*. Unpublished Ph.D. dissertation, University of California, Berkeley, 1958.
- Kitano, H. L. "Validity of the children's manifest anxiety scale and the modified revised California scale." *Child Development*, 1960, 31, 67-72.
- National Film Board of Canada. *Shyness* (23 minutes, sound, black and white film), 1952.
- Ojemann, R. H. "Integrated plan for education in human relations and mental health." *Journal of the National Association of Deans of Women*, 1953, 16, 101-108.
- Rogers, C. R. "The necessity and sufficient conditions of therapeutic personality change." *Journal of Consulting Psychology*, 1957, 21, 95-103.
- Rosenthal, S. "A fifth grade classroom experiment in fostering mental health." *Journal of Child Psychiatry*, 1952, 2, 302-329.
- Ross, Helen. "When in trouble, talk it out." *Just like me, teacher's edition*. Chicago: Scott-Foresman, 1957, pp. 77.
- Rynerson, M. N. *A comparison of the performance of normal and disturbed children on the children's manifest anxiety scale*. Unpublished master's thesis, University of Chicago, 1957.
- Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R. and Ruebush, B. K. *Anxiety in elementary school children: a report of research*. New York; John Wiley, 1960.
- Saskatchewan Department of Education. *Elementary school curriculum guide for health and physical education*. Regina: Queen's Printer, 1956.
- Seeley, J. R. "The Forest Hill Village Experiment." *Understanding the Child*, 1959, 23, 104-110.
- Shacter, H. and Bauer, W. W. *Guidebook for the health and personal development series, Book 4; The girl next door*. Chicago: Scott Foresman, 1948.
- Webster, H., Sanford, N. R. and Freedman, M. "A new instrument for studying authoritarianism in personality." *The Journal of Psychology*, 1955, 40, 73-84.



RELATIONSHIP BETWEEN ACHIEVEMENT SCORES  
AND SELF-ESTIMATES

S. B. KAKKAR  
*Government Training College For Teachers  
Jullundur, India*

Achievement in subjects depends upon the knowledge and information within an individual's conscious mind. It does not tap the unconscious. Thus the individual is well aware of how well he has fared in the subject. Some relationship would be expected between his achievement scores as awarded by the examiner and an independent self-estimate of his performance. To prove this HYPOTHESIS and to determine the magnitude of such relationship the present study was undertaken.

Method

An essay type three-hours' question paper containing 15 questions of Educational Psychology syllabus that they had covered in the classroom was given to 48 teacher-trainees (23 boys and 25 girls) with a note to attempt any five questions; each question carried 18 marks. Shortly afterwards the scripts were evaluated by the investigator (their lecturer in Psychology) and the awards (achievement scores) kept away. No marks were put on the scripts. A week later the investigator told the trainees his own answers to each of the 15 questions with instructions that they should evaluate their answers keeping in view the answers which were expected of them. The independent self-estimates of their performances were thus collected. Efforts were made to control other factors which could affect the trainees' self-estimates. To do this the investigator read every word of the answers and awarded marks on the basis of the answers he had in his mind, and the trainees were required to put their own award on an answer immediately after they heard the investigator's corresponding answer. The only thing kept fluid, as far as possible, was the evaluative ability.

TABLE I  
MEANS AND STANDARD DEVIATIONS OF THE SAMPLE

Scores	N	Means	Standard Deviations
Achievement Scores .....	48	40.6	8.9
Self-estimates .....	48	41.4	7.1



Means and standard deviations of the achievement scores and self-estimates were calculated and the correlation coefficients for the two evaluations computed, using the Pearson product-moment method.

**TABLE II**  
**MEANS AND STANDARD DEVIATIONS AND MEAN DIFFERENCES, FOR BOYS AND GIRLS IN THE SAMPLE**

Sample	N	Means		Standard Deviations	
		Achievement scores	Self-estimates	Achievement scores	Self-estimates
Boys .....	23	36.9	39.6	9.1	7.4
Girls .....	25	44.5	43.2	8.7	6.8
Difference .....		— 7.4	— 3.6		
Significance .....		.001	.001		

**TABLE III**  
**CORRELATION COEFFICIENTS (r) AND STANDARD ERROR (SE)**  
**(CALCULATED THROUGH CLASSICAL FORMULA)**

	N	r	SE
Between achievement scores and self-estimates in general .....	48	.75	$\pm 0.6$
Between boys' achievement scores and self-estimates .....	23	.80	$\pm .007$
Between girls' achievement scores and self-estimates .....	25	.62	$\pm .12$

### Interpretation

1. Of the 48 students 16 were found above average, 13 average, and 19 below average in their achievement.

2. Considering the mean of self-estimates, 16 were found above average, 15 average, and 17 below average.

3. The range of achievement scores is 34, while that of self-estimates is 28, showing that the dispersion of achievement scores is larger than that of self-estimates.

4. The achievement scores deviate more from their mean than the self-estimates from their mean. The two evaluations are thus patterned differently.

5. On the whole the trainees over-estimate themselves though the magnitude of this over-estimation is not much.

6. Girls on the average score considerably more than the boys in their achievement, as well as in their self-estimates.

7. Boys on the average over-estimate their performances, while the girls under-estimate their performances a little.

8. Whereas the girls on the average score above the means of the sample both in their achievements and self-estimates, the boys generally score below the said means both in their achievements and self-estimates.

9. Girls' scores both in achievement and self-estimate deviate less from their means than boys' scores from their means.

10. The correlation coefficient between achievement scores and self-estimates of the trainees is as high as .75 with a standard error of -.06, while the  $r$ 's between achievement scores and self-estimates of boys and between those of girls are .80 and .62 respectively. Boys have been able to judge their performance better and more accurately than the girls.

### Conclusions

There is a significant relationship between achievement scores and self-estimates, this relationship being somewhat greater in the case of boys than in that of girls. This also reflects the fair validity of essay-type examinations, which are often decried, in relation to self-estimates. Students may not be poor judges of their own performance provided they know what is expected of them.

And if they know before the examination what kind of answers are expected of them, they may come up to the instructor's expectations in the matter of their subject-achievement. This fact warrants the need of discussing in the class specific questions on a topic after it has been taught. The lecture method needs to be supplemented by discussion and assignment methods to give students an experience of handling questions and applying the theoretical concepts in everyday use.

## A BIBLIOGRAPHIC SURVEY OF LITERATURE CONCERNING THE PLACE OF GRAMMAR

HOWARD G. AMBURY

That pupils do not use their native language with sufficient competence is stated frequently by both critics of education and by educators themselves. How pupils are to be brought to obtain this competence is a question about which there is much less agreement. One solution frequently proposed is grammar—or some particular kind of grammar. What is the place of grammar in the teaching of English language in the schools is a question that has evoked a vast literature. It is a question as old as Plato and as current as the latest issue of the *English Journal*.

### Formal Grammar and the Traditional Approach

A study of the history of grammar reveals that while grammar was a subject of study in Greece and Rome, it was studied more for its own sake and for literary criticism than for any supposed powers to improve the students' use of their language. It is not clear at what point in the development of Western culture grammar became attached to language teaching. Priscian, in the early part of the sixth century, A.D., used grammar to teach Latin at Constantinople. In the Middle Ages—when the knowledge of Latin among the Church scholars in Europe had become hazy—Latin accidentence was called upon to improve this knowledge of Latin. Later, when English became established as a subject of study in the schools, grammar was included in this study. Since the English grammar texts were based on earlier Latin grammars, much of what they included was Latin grammar. Dykema notes that, "Grammar (English grammar based on Latin accidentence) became so integral a part of Western educational practice that a faith in it was acquired with the education itself." (18:8) It was during the eighteenth century that the theoretical grammarian became a militant force in the regulation of the English language. According to Baugh:

The grammarians (of the eighteenth century) aimed to do three things: (1) to codify the principles of the language and reduce it to rule; (2) to settle disputed points and decide cases of divided usage; and (3) to point out common errors, or what were supposed to be errors, and thus correct and improve the language. All three of these aims were pursued concurrently. But the grammarian set up as law-giver as well. He was not content to record fact; he pronounced judgement. (3: 341)

Robert Lowth and William Ward, grammarians of this type, published grammars in 1762 and 1765 respectively which have been models until quite recent times. Hatfield comments as follows:



(In) some of the composition books of 1910 . . . the treatment of grammar was exhaustive, the organization was strictly on the Latin model . . . Language constructions were either right or wrong, very often depending upon the personal taste (*alias whim*) of the author of the textbook in use. (28: 5)

Sawicki., reporting on the teaching of English language in Alberta from 1905 to 1955, found a traditional grammar dominating the curriculum as late as 1922. He says:

. . . grammar and rhetoric were taught in a formal manner, as a separate subject from composition . . . justified on the grounds that it gave excellent training in logical reasoning . . . and built moral fibre. (48: 144)

The present status of formal grammar in language teaching is dealt with in the next section.

Although the question of what traditional grammar is cannot be readily answered with any single definition, the following list of some of the attitudes attributable to traditional grammarians may serve to delimit the term:

1. A large proportion of students' time is occupied in learning rules and definitions;
2. These rules and definitions are based partly on Latin grammar, partly on the intuition of the grammarians, and partly on the authority of earlier grammarians such as Joseph Priestly.
3. These rules are believed to have some universality of application. That is, language which accords with the rules is 'correct', while language which does not so accord is 'incorrect' at all times and in all cases.
4. Students apply their knowledge of the rules and definitions to the analysis—parsing or diagramming—of sentences and to the correcting of errors in sentences contrived by someone else.
5. The method of teaching is authoriatative and deductive, rather than pragmatic and inductive.
6. Grammar is felt to be a form of discipline for the mind, training the mind to think logically. Logical thinking learned in grammar exercises is expected to transfer to other areas of thought.
7. Grammar, composition and literature are thought to be three separate and distinct studies.

This kind of grammar has been attacked by very many writers using many different arguments. These objections may be summarized as follows: (1) Formal grammar is not a true description of the English language; (2) To analyse the structure on the basis of meaning is to attack the problem the wrong way round; (3) The transfer from knowledge of grammar to the improvement of language competence has not taken place; (4) The structure of English cannot be described in terms of a highly inflected language such as

TABLE I  
A SYNOPSIS OF SEVERAL INVESTIGATIONS RELATING TO GRAMMAR AND COMPOSITION

Investigator and Date	Ref.	Place	Nature of the Study	General Findings
Hoyt, 1906	30	Indianapolis	Compared knowledge of grammar with ability in composition and ability to interpret poem	All correlations were less than 0.30.
Rapeer, 1913	45	Minneapolis	Repeated Hoyt's experiment.	Found similar low correlations.
Asker, 1923	2	Washington	Compared grammar knowledge with sentence sense and composition ability.	Corr. of grammar with sentence sense =0.23; with composition ability =0.37. Both low.
Symonds, 1930	52		Classes taught by five methods and by a combination of all methods.	Combination superior to any single method.
Chalmers, 1941	11	Alberta	Compared knowledge of grammar terminology, grammar analysis skill, extent of vocabulary, composition ability, and prose appreciation.	Only significant relationship was between grammatical analysis and ability in prose composition.
Edmiston and Gingerich, 1942	19		Compared knowledge of usage rules with ability in prose composition.	Knowledge of usage rules does not bear a high relationship with composition ability.
Butterfield, 1945	8	Iowa	Taught matched groups punctuation by functional grammar and by thinking approach.	Those taught grammar learned more grammar but did not improve ability to punctuate.
Silvy Kraus, 1957	32	Oregon	Classes taught by three different methods.	Teaching sentence structure in relation to actual writing practice was most effective.
Robinson, 1959	47	England	Compared marks on grammar tests with marks on essays.	Knowledge of grammar does not significantly improve composition ability.

Latin; (5) The structure of English is not logical but natural: thus formal English grammar cannot train the mind in logical thinking; (6) Children's errors are too varied and complex to be eradicated by rules; (7) Too great emphasis on correctness inhibits the free expression of pupils; (8) The traditional analysis of sentences is an ineffectual waste of the time of pupils and teachers alike. The present writer agrees with Dora V. Smith when she says:

Obviously, the only scholarly attitude to be taken on this question at the moment is one of serious questioning concerning the efficiency of grammatical knowledge in improving speech and writing. (49: 646)

### *Evidence Concerning the Relationship of Grammar to Teaching English*

In the previous section were presented some of the bases on which the teaching of formal grammar has been attacked. In this section the evidence that has been found to support some of these charges is examined.

#### *Formal Grammar in Relation to Communication Skills*

A number of investigators have measured the amount of knowledge of formal grammar possessed by pupils and have compared these results with the results obtained by the same pupils on a measure of ability in composition. Several of these investigations are summarized in Table I. These investigations tend to show that formal grammar as traditionally presented has little effect on students' ability in written composition.

Another group of investigators—including Frogner (23)/1933/, LaBrant (33)/1933/, Godwin in Calgary (24)/1955/, and Gray in Grande Prairie (25)/1959/—compared the written composition of pupils in various grades. Taken together, these four studies seem to indicate that as pupils progress through the grades they improve in both the maturity and correctness of their sentence structure. This may be due to: (a) maturation; (b) practice in writing; (c) the effectiveness of the schools' programs of instruction.

Dressel, Schmidt, and Kincaid (17)/1952/ at Michigan State University and Buxton (9)/1958/ at Alberta conducted experiments which tend to show that writing practice alone does not bring about significant improvement in written composition ability. Buxton found that when students discussed and revised their essays in class they significantly improved their scores on the essay tests without detracting from their scores in other subject areas.

Other investigations done by J. D. Clark (14)/1934/, Briggs (7)/1913/, Benfer (4)/1935/, Boraas (5)/1917/, Price (44)-1923, and Catherwood (10)/1932/ tend to support the contention that



formal grammar as it has been taught has failed to improve significantly the English language competence of pupils.

*Grammar in Relation to Other Purposes  
For Which It Is Proposed*

1. *In relation to the teaching of foreign languages.* Rivlin (46) /1934/ found that teachers of English language tended to rate many more items of grammar as important to the teaching of foreign languages than did the teachers of these languages themselves. Of the 187 items of grammar tested, only 23 resulted in agreement among the teachers as to their importance. Rivlin agrees with Kaulfers (31), Pooley (42), and Pamela Gradon (20) that English language teachers should teach only those items of grammar that are important to a study of English and that the grammar that is needed for the teaching of foreign languages should be taught in the foreign-language classes. Parker /1941/ at Alberta found that "Stenography II and French II students are greatly handicapped by the fact that they have not mastered formal English grammar." (40: 210) However, certain weaknesses in this investigation cast some doubt on the reliability of this conclusion.

2. *In relation to literature appreciation and reading ability.* Chalmers (11) found little relationship between the ability to appreciate literature and knowledge of grammar terminology. Strom /1955/ found that there was "little, if any, relationship between the pupils' comprehension of ten selected passages of poetry and literary prose and their ability to classify crucial elements of grammar and syntax in the sentences of these passages . . ." (51: 131) No evidence was found to support the hypothesis that the study of grammar is necessary for the improvement of reading or for the appreciation of literature.

3. *In relation to college and university entrance.* Although one reason frequently advanced to justify formal grammar teaching is that the colleges demand that students know grammar, studies made by Smith and McCullough (50) /1935/, Litsey (35) /1956/ and the N.C.T.E. Committee on High School-College Articulation /1961/ fail to support this view. The report of the N.C.T.E. committee concludes:

The goal in grammar seems to be enough terminology to make the college instructor's theme corrections intelligible to the students . . . In all of these language matters, the college statements ask for a functional knowledge of the material.. To recognize but not be able to use does the student little good. (38: 411)

4. *In relation to the requirements of the Alberta Department of Education.* To determine the grammar requirements of the Department of Education the writer first examined the textbooks author-

ized for use in the English language programs of the junior and senior high schools and the curriculum guides supplied for these programs and, second, he analyzed the examinations set by the Department for grade nine and grade twelve.

The curriculum guides for the various English language programs tend to stress a functional approach to grammar and usage and to stress usage over formal grammar as a guide to what is good language. The textbooks that have been authorized tend to support this philosophy of language teaching. There is, however, a book called *Supplementary Exercises in Grammar*<sup>1</sup> for use in grade nine, and there are numerous exercises in *Creative Composition*<sup>2</sup> for grade ten, and *Guide to Modern English*,<sup>3</sup> a supplementary text for grades eleven and twelve.

To determine the grammar content of the Departmental examinations, the June 1960 examinations in Social Studies-Language B for grade IX and the A and B papers in English 30 were analysed. Of the 213 possible marks on the grade nine paper, 92 could be distinguished as marks for grammar and usage; another 25 marks could be lost for mechanical errors in the essay. Of the one thousand marks available on the two grade twelve papers, 80 could be classified as marks for grammar and usage, while an additional 125 marks could be lost for mechanical errors on the essay. There were no questions in either examination requiring students to analyse sentences or to define grammatical terms. Clearly, although the Department expects students to know a certain amount of grammar and to be familiar with current English usage, they do not require an extensive course in formal grammar.

### *Formal Grammar as a Description of the English Language*

Any literate person who compares his own use of English with the prescriptions of many recent grammars will come to the conclusion that there is something radically wrong with either his own use of English or with the grammars. A further comparison of the rules of the grammars with the findings of such usage researchers as S. A. Leonard (34), Marchwardt and Walcott (36), and C. G. Fries (21) will confirm his conclusion that formal grammar is not an accurate description of the actual usage of present-day speakers of English.

No investigations were found which specifically sought to determine the extent to which formal grammar is a true description of

1. Province of Alberta, Department of Education, *Supplementary Exercises in Grammar for Grade IX*. Edmonton: Queen's Printer, 1958.
2. R. J. and W. C. McMaster, *Creative Composition*. Toronto: Longmans-Green, 1957.
3. Richard K. Corbin, Porter G. Perrin, and Earl W. Buxton, *Guide to Modern English*. Toronto: Gage. (No date given.)



our language, but the following objections put forth by the usage advocates seem to have some validity: (1) languages change much faster than the rules change; (2) efforts in the past—the French Academy, for example—to fix language patterns have signally failed; and (3) languages change, but this change is not *decay*. The usageists suggest that teachers be sure that those usages they condemn in their students' work are those about which there is no doubt.

### *Formal Grammar as a Current Educational Issue*

Some evidence was found that formal grammar is still extensively taught and advocated and that the question concerning the place of grammar in the English language program is a current one. For example, Parker notes that, "Many teachers consulted . . . stated that they used their own texts and taught considerably more formal grammar than was suggested in the outline." (40: 5) The traditional attitude toward grammar as discipline is revealed in an editorial from the *Vancouver Province* quoted in the *Edmonton Journal* of April 23, 1959. The following is an excerpt:

Grammar is not an exciting subject . . . but it trains young minds to think, it forces them to learn and abide by rules. . . . To scrap grammar would be to scrap the philosophy that learning is a disciplining process.

A teacher from Saskatchewan has described to this writer a lesson presented by the supervising teacher of the Lloydminster School Unit in 1959. The lesson in formal grammar, given to a class of pupils from grade four to grade eight, lasted for more than two hours. Hall (27), Anderson (1), Bossone (6), Pooley (43) and Zahner (54) agree that much of what is taught in the schools is still formal grammar. Day (16)/1951/, from personal visitations at fifty-three different Iowa high schools, concluded that the content and organization of English language courses was determined by the textbook and that grammar and usage were most frequently taught as separate materials from actual speaking and writing situations. Mary Wood Dawson compared the presentation of pronouns in 23 school grammars used during the 1940's with that in earlier grammars and concluded, "Modern books, like those of the past, drill students constantly on names, definitions and rules." (15: 36) Patrick Groff (26) studied several well-known textbooks in English language for the junior and senior high-school grades published between 1955 and 1960. He found that the average of the percentages of their volumes devoted to parts of speech was twenty-eight per cent. In other words, these modern textbooks imply that almost 30% of English language teaching time should be devoted to the parts of speech alone.

A survey of the literature with regard to the relationship between grammar and the teaching of English language seems to indicate



that: (1) formal grammar taught in isolation from its actual use in communication is of doubtful value in improving communication skills; (2) the value of teaching extensive formal grammar for purposes other than communication skill is not established; (3) formal grammar is not an adequate description of the English language; but (4) none-the-less, much formal grammar is still advocated and taught and it is included in contemporary textbooks..

### The Alternatives

Critics of the traditional approach to grammar have proposed a wide variety of alternatives. For purposes of simplicity, these are grouped into four 'approaches': (A) The Writing, or Thought, Approach; (B) The Functional Grammar Approach; (C) The Usage-Semantics Approach; and (D) The Scientific Grammar Approach. This grouping of the alternatives is arbitrary and a matter of convenience rather than of fact, since such clear distinctions as are hereby implied do not, in fact, exist. Similarly, though there is the danger of over-simplification in the following generalizations, some attitudes that appear to be common among advocates of each approach are listed. This again is a matter of convenience rather than of fact.

#### *The Writing, or Thought, Approach*

The advocates of this approach would tend to agree that: (1) Writing, speaking, reading, listening and social competence are inseparable skills that are best learned together; (2) Pupils will learn to write correctly and effectively when they have something to say that is important to themselves; (3) Writing is learned by practice in writing, rather than by artificial drill in grammar; (4) Experimental evidence shows that the writing approach is effective in improving communication skills; (5) The grammar that is to be taught is to be determined by the actual mistakes that the pupils make in their own writing; and (6) Those elements of grammar that it is found must be taught will best be taught in close conjunction with actual writing practice.

#### *The Functional Grammar Approach*

Summarizing the functional grammar approach to the teaching of English language, one might list the following points on which functionalists generally agree: (1) The usage of current, reputable resources is generally to be preferred to rules of grammar in determining what is right and what is wrong; (2) Certain principles and rules of grammar must be taught; (3) A direct, systematic attack must be made to ensure the mastery of both grammar and usage items; (4) The specific items of grammar and usage to be taught are

determined by diagnostic testing and by error counts to determine which items cause difficulty; (5) The attack should be centered on one or two items at a time until these have been mastered; (6) Much writing is to be done by the students and the grammar they have learned is expected to *function* in making their sentences correct and forceful and their paragraphs unified, coherent, and emphatic. Grammar and writing are to be closely integrated; and (7) The teacher dominates the classroom situation. She has prepared a lesson plan covering the five learning steps; she will test the pupils to determine which items to stress; and she will plan the drill lessons to ensure mastery of these items.

### *Usage and Semantics as Functional Approaches*

Many educators feel that formal grammar is imperfect as a description of the actual language, and therefore, not a reliable guide to correct speech and writing. For this reason they prefer to teach that grammar which is a description of the actual language used by reputable speakers and writers. For the sake of simplicity, this type of grammar is often called *usage*. The usage approach is not properly a different approach to the teaching of grammar but, rather, a different approach to what constitutes 'correctness'. Fries writes, "The only grammatical correctness there can be in English . . . must rest on usage . . . There can thus never be . . . an error that is both very bad and very common." (22: 33) Most usagists classify language usage into the following, or similar, levels: non-standard, colloquial, informal spoken, informal written, formal spoken, formal written, and literary. They advocate that the informal written and the formal spoken and written levels be taught in the schools since these are the levels that the student may not learn by himself, and they are the levels that will be of use to him in the adult world of successful people. Porter Perrin calls it the language "used by educated people in carrying out their affairs." (41: iii)

The science of meanings, called *semantics*, deals with insights into the ways words acquire different shades of meaning when used in different contexts. Stuart Chase (12) and S. I. Hayakawa (29) have shown that words can be used to reveal fine shades of meaning, but that words can also be used by unscrupulous people to *hide* meaning. An understanding of the ways words work should enable students to use language for more effective communication.

### *Scientific Grammars*

1. *Structural linguistics*... Linguistics is the scientific study of languages. Muinzer (37) identifies three specialities in linguistic study: historical linguistics—studies the ways languages have changed to become what they now are; comparative linguistics—compares dif-



ferent languages to trace the development of one language from another; and descriptive linguistics—attempts to determine and classify the distinguishing features of a given language at a given time. A part of descriptive linguistics—often called structural linguistics—is concerned with the interlocking system of contrastive patterns in a language which signal grammatical meaning. These are patterns of sound, patterns of word forms, and patterns of word order. Structural linguistics is composed of two major branches: phonology—concerned with the sounds of a language; and grammar—concerned with how these sounds are grouped into meaningful utterances. Sometimes *structural grammar* refers to both these branches and sometimes to the latter only.

One of the distinctive ways in which structural grammar differs from traditional grammar is in the basis of its analysis. Newsome describes the structural basis of analysis as follows:

Since parts of speech are arbitrary classifications in any grammar, it is not surprising that individual linguists select slightly different criteria in determining these form classes. It is possible to classify parts of speech primarily on the basis of inflectional endings as Trager and Smith do, or on the basis of syntactic functions, as Long suggests . . . Or the criteria may include both form and word order—a method employed by Fries, Roberts, Francis, and Lloyd and Warfel . . . But despite differences, these systems of classification have one important characteristic in common: they all work inductively and arrive at operational definitions which can be verified; they do not work deductively from semantic or functional definitions.

In a classification of words based upon form and word order, five signals function: inflections and derivational suffixes (special formal characteristics of the words themselves); word order; structure words, or function words, which accompany the words; and, occasionally, stress. (39: 4)

The structuralists criticize traditional grammar for basing the analysis of sentences partly on function and partly on meaning. For example, all the *grammatical* information about the parts of the following sentence may be ascertained, even though we do not know its lexical meaning: 'A ticky tove rimpls the most dinsy moosefores snavely.' Structuralists agree that there are a limited number of basic sentence patterns and an infinite number of variations of these patterns. While the relationship is not exact, intonation and stress often signal the type of punctuation required and the completeness or incompleteness of a sentence. Exercises in structural grammar tend to be synthetic, rather than analytic.

2. *Generative, or transformational, grammar.* A recent development in the field of structural synthesis is propounded by Noam Chomsky in his *Syntactic Structures*. (13) The basic concept of Chomsky's grammar is the 'kernel' sentence. This is a simple, active, declarative sentence composed of a noun phrase and a verb



phrase which have been derived from his 'phrase structure' rules. "The boy hits the dog." is a kernel sentence. From the kernel sentences are *generated* all other sentences by means of *transformations*. A transformation is a rule or formula that introduces a new element into a kernel sentence, or rearranges the parts of a sentence, or both. For example, by one transformation the adjective 'little' can be added to the sample kernel sentence in only two places—before boy or before dog—to generate the sentences: "The little boy hits the dog." or "The boy hits the little dog." or "The little boy hits the little dog." By another transformation any of these can be made passive: "The little dog is hit by the little boy." The only 'ungrammatical' sentences are those that cannot be generated by the rules. "The boy little dog the hits." is ungrammatical because no series of transformations will produce it. Owen Thomas (53) gives a description of the system which, while necessarily oversimplified, will give the reader an idea of the nature of the system. The *English Journal*, May, 1963, contains several articles concerning generative grammar.

### Conclusions

While there is considerable evidence to show that formal grammar as traditionally taught has failed to produce the expected improvement in students' language competence and while there is little evidence favoring extensive knowledge of formal grammar for any of several other purposes for which it is commonly advocated, yet formal grammar is still advocated by some authorities and still taught by some teachers. In spite of the doubts about the value of formal grammar, there seems to be a felt need for some tangible content in the English language program. Perhaps a faith in formal grammar is instilled in the pupils along with the grammar itself. Some evidence shows that it would be unwise to depend on maturation or on writing practice alone to improve students' communication skills.

Four groups of alternatives to formal grammar as traditionally taught have been identified in the literature: the writing, or thought approach; the functional grammar approach; the usage-semantics approach; and the scientific grammar approach. Each of these approaches has some merits as well as some weaknesses. Since there is no one 'right' way to teach language, teachers should read widely in the literature and be prepared to adopt new and better ways of presenting their subject.

### REFERENCES

- Anderson, Wallace L., "Structural Linguistics: Some Implications and Applications," *English Journal*, 46 (October, 1957) 410-18.
- Asker, William, "Does Knowledge of Grammar Function?" *School and Society*, 17 (January 27, 1923) 109-111.

- Baugh, Albert C., *A History of the English Language*. New York: Appleton-Century-Crofts, 1935.
- Benfer, Mabel C., "Sentence Sense in Relation to Subject and Predicate," M.A. thesis, University of Iowa, 1935.
- Boraas, Julius, "Formal English Grammar and the Practical Mastery of English." Ph.D. thesis, University of Minnesota, 1917.
- Bossone, Richard M., "Let's Talk Sense About English," *English Journal*, 43 (October, 1954) 371-74.
- Briggs, T. H., "Formal English Grammar as Discipline," *Teachers College Record*, 14 (1913) 251-343.
- Butterfield, Claire J., "The Effect of a Knowledge of Certain Grammatical Elements on the Acquisition and Retention of Punctuation Skills," Ph.D. dissertation, University of Iowa, 1945.
- Buxton, Earl W., "An Experiment to Test the Effects of Writing Frequency and Guided Practice upon Students' Skill in Written Expression," *Alberta Journal of Educational Research*, V (June, 1959) 81-99.
- Catherwood, Catherine, "Relationship Between a Knowledge of Rules and Ability to Correct Grammatical Errors." M.A. thesis, University of Minnesota, 1932.
- Chalmers, John West, "A Study of the Relationship Between Ability in Formal Grammar and Ability in Literary and Linguistic Fields." M.A. thesis, University of Alberta, 1941.
- Chase, Stuart, *The Power of Words*. New York: Harcourt Brace, 1953.
- Chomsky, Noam, *Syntactic Structures*. The Hague, Netherlands: Mouton and Co., 1957. (Available from Humanities Press, Inc., 303 Park Avenue, South, New York 10, N.Y.).
- Clarke, J. D., "A Four-Year Study of Freshman English," *English Journal*, (Col. Ed.) 24 (May, 1935) 403-10.
- Dawson Mary Wood, "The Passing of the Pronoun," *English Journal*, 45 (January, 1956) 34-37.

- Day, James Edward, "The Teaching of English in Iowa High Schools." Doctoral dissertation, University of Iowa, 1955.
- Dressel, Paul, John Schmidt, and Gerald Kincaid, "The Effect of Writing Frequency upon Essay-type Writing Proficiency at the College Level," *Journal of Educational Research*, 44 (December, 1952) 285-91.
- Dykema, Karl W., "Historical Development of the Concept of Grammatical Properties." *Applied English Linguistics* (Harold B. Allen, Ed.) New York: Appleton-Century-Crofts, 1958.
- Edmiston, Robert W. and C. N. Gingerich, "Relation of Factors of English Usage to Composition," *Journal of Educational Research*, 36 (December, 1942) 269-71.
- English Association Symposium, (Vivian de Sola Pinto, Ed.) *The Teaching of English in Schools*. London: Macmillan, 1948.
- Fries, C. C., *American English Grammar*. English Monograph No. 10 of the N.C.T.E. New York: Appleton-Century-Crofts, 1940.
- , *The Teaching of English*. Ann Arbor, Michigan: George Whar Publishing Co., 1949.
- Frogner, Ellen, "Problems of Sentence Structure in Pupils' Themes," *English Journal*, 22 (November, 1933) 742-49.
- Godwin, Lois Ruth, "An Analysis of the Sentence Structure and Paragraph Writing of City of Calgary Students in Grades V, VIII, and IX." M.Ed. thesis, University of Alberta, 1955.
- Gray, Mary E., "An Analysis of Language Themes in Grade Five, Grade Eight, and Grade eleven." M.Ed. thesis, University of Alberta, 1959.
- Groff Patrick J. "Is Knowledge of Parts of Speech Necessary?" *English Journal* 50 (September, 1966) 413-15.
- Hall, Robert A., (Jr.) *Leave Your Language Alone*. Ithaca, N.Y.: Linguistica, 1950.
- Hatfield, W. W. "Progress in Teaching English," *N.A.S.S.P. Bulletin*, 136 (February, 1963) 5-10.
- Hayakawa, S. I., *Language in Thought and Action*. New York: Harcourt Brace, (Rev.) 1949.
- Hoyt, Franklin S., "The Place of Grammar in the Elementary School Curriculum," *Journal of Educational Psychology*, 4 (March, 1913) 125-37.
- Kaulfers, Walter V., *Four Studies in Teaching Grammar*. Stanford University, California: University Bookstore, 1945.
- Kraus, Silvy, "A Comparison of Three Methods of Teaching Sentence Structure," *English Journal*, 46 (May, 1957) 275-81.
- LaBrant Lou, *Study of Certain Language Developments in Children in Grades IV to XII*. Worcester: Clark University Press, 1933.
- Leonard, Sterling A., *Current English Usage*. N.C.T.E. Monograph No. 1. Chicago: N.C.T.E., 1932
- Litsey, David M., "Trends in College Placement Tests in Freshman English," *English Journal*, 45 (May, 1956) 250-56.
- Marckwardt, Albert H. and Fred Walcott, *Facts About Current English Usage*. N.C.T.E. Monograph No. 7. New York: Appleton-Century-Crofts, 1938.
- Muinzer, Louis A., "Historical Linguistics in the Classroom," *Illinois English Bulletin*, 48 (January, 1961) (Special republication) 24-58.
- National Council Teachers of English, "What the Colleges Expect," *English Journal*, 50 (September, 1961) 402-12.
- Newsome, Verna L., *Structural Grammar in the Classroom*. Oshkosh, Wis.: Wisconsin Council of Teachers of English, 1961.
- Parker, Patricia E., "A Comparison of Student Efficiency in English Grammar with that in Two Other Selected Subjects of the High School Course." B.Ed. dis. University of Alberta, 1941.



- Perrin, Porter G., *Writer's Guide and Index to English*. New York: Scott-Foresman (Revised) 1950.
- Pooley, Robert C., "Forever Grammar," *N.A.S.S.P. Bulletin*, 136 (February, 1946) 45-49.
- "Looking Ahead in Grammar," *N.A.S.S.P. Bulletin*, 39 (September, 1955) 56-61.
- Price, Edwin A., "A Study in the Evaluation, Frequency, and Method in the Elimination of Sentence Errors among Junior High School Students." Master's thesis, Leland Stanford University, 1923.
- Rapeer, Louis W., "The Problem of Formal Grammar in Elementary Education," *Journal of Educational Psychology*, 4 (March 1913) 125-37.
- Rivlin, Harry N., "English Grammar as Preparation for the Study of a Modern Language," *English Journal*, 23 (March, 1934) 202-07.
- Robinson, Nora, "The Relation Between Knowledge of English Grammar and Ability in English Composition" *British Journal of Educational Psychology*, 30 (June, 1960) 184-86.
- Sawicki, Stanley W., "The Development of the English Program in the Secondary Schools of Alberta," *Alberta Journal of Educational Research*, IV (September, 1958) 142-51.
- Smith, Dora V., "English Grammar Again," *English Journal*, 27 (October 1938) 643-49.
- Smith Dora V. and Constance McCullough, "An Analysis of the Content of Placement Tests in Freshman English Used by 130 Colleges and Universities," *English Journal*, 25 (January, 1936) 17-25.
- Strom, Ingrid M., "Does Knowledge of Grammar Improve Reading?" *English Journal*, 45 (March, 1956) 129-33.
- Symonds, Percival M., "Practice versus Grammar in the Learning of Correct English Usage," *Journal of Educational Psychology*, 22 (February, 1931) 81-95.
- Thomas, Owen, "Generative Grammar: Toward Unification and Simplification," *English Journal*, 51 (February, 1962) 94-99.
- Zahner, Louis, "The Teaching of Language," *English Journal* 44 (November, 1955) 443-50.

# YESTERDAY'S INFANT SCHOOL

K. H. THOMSON

*Faculty of Education, Edmonton*

## Part I

There would appear to be a steadily increasing interest in earlier entrance to institutions of formal education. Until recently one of the main manifestations of this interest has been the increasing number of parents who express the desire to have children enter grade one at age five or younger. Now we are seeing not only a continuation of this desire but also more interest and action in kindergarten organization. In 1962 this movement received renewed official recognition through the promulgation of Department of Education regulations which reinforced the permissive legislation of the Alberta School Act (Sections 395 and 396) and which set out the conditions under which a kindergarten might operate. These regulations make it necessary to obtain ministerial approval before a pre-school institution can call itself a kindergarten. If there is not this approval the institution is a play school or nursery school and comes directly under the regulations of the Department of Welfare rather than those of the Department of Education. Without this approval such a pre-school institution may not advertise itself as a kindergarten. Factors taken into consideration when an application is being appraised are the nature of the accommodation, the size and qualifications of the staff, and the nature of the program and equipment proposed. With respect to staff qualifications it is expected that newly appointed kindergarten teachers will hold a teaching certificate and have included kindergarten courses in their training. It is of interest to note that these institutions are subject to Department of Education inspection, that approval is given on a year to year basis, and that approval may be withdrawn at any time within the year. According to the Annual Report of the Department of Education for 1963 fifty-four applications for private kindergartens were approved. At the time of writing about two hundred applications have been received by the Department of Education and approximately one hundred fifty of these have been approved. Further departmental interest in this level of instruction is shown by the development of a guide to a kindergarten program of studies. The University of Alberta also shows recognition of the trend by the development of a teacher training course called Curriculum and Instruction in the Kindergarten.

Although it is evident that there is considerable activity in the area of private kindergartens very little is being done by regularly

constituted school boards. The West Jasper Place School District did operate one or two kindergartens and the City of Grande Prairie has made a start in this direction. The school systems associated with the Department of National Defence also operate kindergarten classes. No doubt the number of school boards entering this field of education will gradually increase but since the principle of fees is recognized for instruction at this level there will be less pressure to move from private to public sponsorship. From the point of view of the grade one teacher it would seem desirable that the trend become very widespread quickly now that both kindergartens and play schools have become so much more common in the urban areas. That the movement will continue to spread there can be little doubt, but the speed with which it will spread is difficult to forecast. One immediate side-effect of the regulative action however, which will tend to clear the way for the growth of the kindergarten movement, is the development of a clear distinction between a play school and a kindergarten.

In view of this interest in early school attendance it would seem appropriate to review the educational ideas and practice of one who promoted an early interest in infant education and infant schools. Such a person was Robert Owen. \*

## Part II

As a backdrop to the educational endeavors of Robert Owen a brief look at Owen himself would seem useful. He was born in Newtown, North Wales, May 14, 1771, the son of a saddler, and, after a successful business career, much public service and much travel, he died in the town of his birth on the seventeenth of November, 1858.

According to his recollection he attended a school established in the mansion of a neighboring estate and his formal education lasted from the age of four or five to the age of nine.

In schools in these small towns, he wrote, it is considered a good education if one could read fluently, write a legible hand, and understand the four first rules of arithmetic. And this I have reason to believe was the extent of Mr. Thickness's qualifications for a schoolmaster, because when I first had acquired these small rudiments of learning, at the age of seven, he applied to my father for permission that I could become his assistant and usher, as from that time I was called while I remained in school. And thence forward my schooling was to be repaid by my ushership. As I remained at school about two years longer, those two years were lost to me, except that I thus early acquired the habit of teaching others what I knew. But at this period I was fond of and had a strong passion for reading everything which fell my way. As I was known to and knew every family in the town, I had the libraries of the clergymen, physician, and lawyer—the learned men of the town—thrown open to me, with permission to take home any volume which I liked, and I made full use of the liberty given to me.<sup>1</sup>



Owen's informal education, through his reading and the wide range of the ideas and backgrounds of the great many personal acquaintances he had, was far more extensive than his formal education. His interest in reading he attributed to a delicate constitution which prevented his engaging in the activities that consumed the time of more fit children. His success in coming to know a great many prominent people from many walks of life he attributed to the quality of the textiles from the factory he managed, "for," as he states, "the best manufacturer in any branch of cotton manufacture became, in those days, a person of public celebrity, and my name now stood prominent before the Manchester public."<sup>2</sup> This reputation caused him to become a member of the Manchester Literary and Philosophical Society and he was on very friendly terms with members of the staff of Manchester College. By 1813, Owen was widely known for his keen interest in social reform and his personal contacts were very widespread. His acquaintances among politicians included the Prime Minister, Lord Liverpool, among churchmen the Archbishop of Canterbury, among political economists such as Malthus, James Mill and Ricardo, among princes the Duke of Kent—father of Queen Victoria, Grand Duke Nicholas of Russia—later to become Tsar Nicholas I, the Duke of Orleans—later to become King Louis Philippe of France. There would seem to be no doubt that Owen was widely known and aroused a great deal of interest during the first quarter of the nineteenth century. As a result of his diversified contacts and interests his influence was fairly extensive.

Owen's travels abroad were quite extensive too, including visits to the United States, Mexico, St. Domingo, Jamaica, France, Switzerland and Germany.

With respect to education, while in Europe he visited the schools of Father Oberlin in Geneva, of Pestalozzi at Yverdon and of Fellenberg at Hotwyl, and it would appear that he never missed an opportunity to tell these people about the principles and operation of his school. In 1813 he devoted much time and effort to the promotion of

Dr. Bell's and Mr. Lancaster's plan for educating the poor, and in other public matters which were then beginning to occupy the attention of benevolent men, for this was at the commencement of the new era for ameliorating the condition of the poor and for educating their children.<sup>3</sup>

Even though these efforts were often poor and misguided he took a keen interest in a great many. As he said: "The beginning of some education, however defective, was much better than the entire neglect of it; and I confidently expected that when once commenced it would gradually progress toward a much more mature state."<sup>4</sup> He claims to have had considerable influence on the de-

velopment of Prussian education, which in turn became so influential elsewhere through the influence of Victor Cousin and others. He tells us that the Prussian Ambassador, Baron Jacobi, requested an introduction to him and sent copies of Owen's writings to his King

Who so much approved of them as to write an autographed letter to me expressing his high approbation of my sentiments on national education and on government, and stating that he had in consequence given instruction to his minister of the interior, to adopt my views on national education to the extent that political condition and locality of Prussia would admit.<sup>5</sup>

Toward the end of the 1820's Owen's influence was waning. His many cooperative projects were short-lived and he lost touch with the main stream of reform. By 1837 he was receiving practically no attention although he continued to be very active with his writing and speaking. At the age of eighty-two he turned to spiritualism as a means of promoting his social and educational views.

### Part III

To round out the background against which Owen's educational theory and practice was developed, a few comments to aid in recalling the state of society in his times might be appropriate. It was a period of fairly rapid transition with all the social dislocation that this involves. The developments of science promoted the enclosure movement with its consequent disruption of agricultural patterns and made possible the factory system which practically destroyed domestic industry and furthered the enclosure movement. These forces created a large uprooted segment of the population which was now without either economic or social stability. The effects of these social changes were also reflected in serious slum conditions developing very rapidly in many urban areas. These conditions would have had grave human consequences if only certain segments of the adult population had been involved, but it was far more serious when it involved a large number of children, many of whom were very young indeed.

Congestion, child labor, child neglect, low wages and long working hours became very common. With respect to the first, one medical officer estimated that in one area of London there were 581 persons living on half an acre. Regarding child labor M. T. Sadler, M.P. said:

Our ancestors could not have believed it possible—posterity will not believe it true—that a generation of Englishmen could exist or had existed, that would work lisping infancy of a few summers old, regardless alike of its smiles or tears, and unmoved by its unresisting weakness, twelve, thirteen, fourteen, sixteen hours a day, and through the weary night also, till, in the dewy morn of existence, the bud of youth was faded and fell e'er it was unfolded.<sup>6</sup>



The problems arising from both parents working where care of children is not provided is not new to us.

No doubt female and child labor did much to force down the level of pay to heads of households because of competition in general and because the scale of pay to females and children was less than that to males. We are told that girls would work for 6 to 8 shillings a week at jobs which had commanded from 18 to as high as 45 shillings for men. The development of the 'truck system' or company store often added to the misery because of uncontrolled prices where wages were paid in goods.

With respect to long working hours, we read of amazing feats of endurance. A train fireman gave evidence that he had worked continuously as long as twenty-nine hours, fifteen minutes, and his week's working time on this particular occasion was eighty-eight hours forty minutes. In June of 1863 a London newspaper carried an item headed: 'Death from Overwork' which told of a Mary Ann Walkley, age twenty, who worked in a milliner shop. Her average daily hours of work were fourteen and a half, but in rush seasons she frequently worked for thirty hours on end. Her death occurred during such a season. With respect to conditions in large factories one inspector reported that in one mill he found women who for many weeks had been employed from six in the morning till midnight with less than two hours off for meals. Proposals to control working hours met with strong resistance and were looked upon as a threat to the economy. According to Marx the working day varied from eight to eighteen hours, often with no regular time for meals being allowed.

No doubt the extremes of hardship received particular attention, but it is certain that the situation was very bad, that a thorough going program of reform was needed, and it was to the cause of reform that Robert Owen devoted most of his energy throughout his long life.

#### Part IV

Elementary schooling for most children in Owen's time was either not available or took the form of Charity Schools or private 'baby-sitting' types of institutions such as Dame Schools and those established by tradesmen in their shops. The following is a description of a tradesman-teacher.

William Woodcock, more familiarly called Billy Bingo—a jolly little man, who united the two vocations of schoolmaster and baker, while his wife boiled toffee and barley sugar for the children. He was a very merry, conceited little fellow, who thought these three pursuits particularly congenial, and often boasted that he nourished both body and mind . . . He used to come whistling or humming a tune, hot out of his bakehouse, with his shirt sleeves rolled up to his shoulders, sit down at once to cool himself and hear our lessons; and then, having set us our copies, he strutted



humming away again to set his bread. A more contented creature never existed: to hear him talk, you would imagine he had educated all the great men in the country.<sup>7</sup>

All such teachers were by no means as kind and jolly as William Woodcock. Smith quotes Joseph Lancaster as saying that such teachers were "but too often the refuse of superior schools, and society at large. Their earnings were small, their schools unhealthy, their drunkenness was 'almost proverbial.' They punished harshly, and regarded the lash as 'the only properly governing medium.'"<sup>8</sup>

Another type of school which represented something of a new trend was the Sunday School. These schools were started mainly by Robert Raikes and had as their main purpose the occupation of the child's time when he was freed from work on Sundays and to occupy this time profitably by inculcating concepts of good conduct. Neglected children were becoming a public menace from the point of view of established society. As the Reverend Charles Moore stated it, the purpose of the Sunday School

Was to furnish opportunities of instruction to the children of the poor without interfering with any weekly industry; to infuse into the tender minds of infancy ideas of decency, sobriety and industry; to inure them to early habits of regularity in their attendance at church and to teach them how to spend the leisure hours of Sunday to their own improvement, advantage and happiness, which are now almost universally consumed in idleness, profanation and riot.<sup>9</sup>

These schools often opened at 8:00 a.m. and included morning church service. After a noon break they generally reassembled at 2:00 p.m. and remained in session until 7:00 p.m. in summer and 6:00 p.m. in winter.

Another new movement was the monitorial type of school introduced by Joseph Lancaster and Andrew Bell to which reference has been made. The schools established by Lancaster were non-denominational, while those established by Bell reflected the concern of the clergy over such secular institutions.

In spite of various types of elementary schools it would seem proper to credit Owen with being the founder of the true infant school in England and certainly among the forerunners of this movement elsewhere.

## Part V

G. D. H. Cole writes:

Robert Owen, shop-boy and manufacturer, factory reformer and educationist, socialist and co-operative pioneer, trade union leader and secularist, founder of ideal communities and practical man of business, was something of a puzzle to his own generation, and is no less a puzzle to posterity. Surely no man ever founded so many movements, and yet had so simple and unvarying a body of ideas.<sup>10</sup>

Basic to Owen's theoretical position were his ideas regarding the nature of religion, the nature of man and man's self-realization. The

key to self-realization, to human happiness, lay in the development of a spirit of universal charity, which, for him, depended on overcoming the intolerance and divisiveness generated by sectarian religion and other forces of disunity among men. He states:

... I am not of your religion, nor of any religion yet taught in the world!—to me they all appear united with much—yes, with very much!—error!<sup>11</sup>

The errors to be found in all sectarian religion made rational behavior impossible. By rational behavior he meant behavior governed by universal charity. Society needed to permit the most unlimited religious freedom. Indeed, society needed to transcend religious doctrine. Universal charity disregards creeds, he argued.

Religious sectarianism was not the only divisive force in society however. Economic insecurity was also a major cause of human disunity. The abolition of this cause of disunity would come through the employment of cooperative enterprise. Through cooperation men could produce more than they needed and thus each would get all the goods and services he needed. This would remove the desire for individual accumulation of wealth, indeed such a desire would appear to be irrational.

And what of the nature of man, which, freed from the blight of creed and covetousness, could operate in accordance with the spirit of universal charity? To begin with: "By my own experience and reflection I had ascertained that human nature is radically good, and is capable of being trained, educated, and become united, good, wise, wealthy and happy."<sup>12</sup>

Although he placed great emphasis on the power of environment he recognized a basic nature which did impose limits on adaptability. But within the limits of the potentiality of each the social environment determined or directed development. With respect to this matter he stressed the lack of personal responsibility for the course of individual development.

The basic quality to be found in man is the desire to obtain happiness which is primary cause of all his actions. In addition he was imbued with the desire to "sustain, enjoy, and propagate life."<sup>13</sup> And he has the potentiality to receive, convey and compare ideas, and the potentiality to become conscious of these mental processes. These mental powers develop with the growth of the individual. The way in which a man's potential develops, which varies from individual to individual in some degree, depends on his experience "from the objects around him, and chiefly from the example and instruction of his immediate predecessors."<sup>14</sup>

It will be apparent that man and society cannot depend on the natural unfolding of human nature to secure happiness. 'Positive'



education is necessary for the proper development of the individual just as 'positive' governmental legislation is necessary for the proper development of society. Proper environmental and directive measures could do wonders to reform the individual and society. "The fundamental principle on which all these essays proceed is," he writes, "that 'children collectively may be taught any sentiments and habits,' or in other words, 'trained to acquire any character.'"<sup>15</sup> And again:

Any general character, from the best to the worst, from the most ignorant to the most enlightened, may be given to any community, even to the world at large, by the application of proper means; which means are to a great extent at the command and under the control of those who have influence in the affairs of men.<sup>16</sup>

Owen saw in the nature of man far reaching implications for education and schooling. In the first place, since so much of man's character was formed for him, it was imperative that there be a uniform national system of education and that this system be universally applied. Education should not be an object of inappropriate parental application nor of parental neglect. And since character formation begins at birth formal education should begin at the earliest possible age in a manner that would provide the least possible opportunity for the acquisition of bad habits and in a manner that would provide greatest opportunity for learning universal charity.

Owen, therefore, would have children enter the infant school as soon as they could walk with a reasonable degree of assurance. At times he suggests age two and at other times age three. In any case, certainly by age three, the little people would enter boarding schools. In the main, they would be away from parental influence but parents would be expected to visit them often during meal times and at other appropriate times. The period of infant school education was to be divided into two levels, the first and the second preparatory schools. The first of these levels we would perhaps call play school and the second we might call kindergarten, although there would probably be some extension into our grade one level in the second preparatory school. At this latter level a beginning in the development of basic skills would take place. The elementary school or third level would extend from age six to ten, although he would recommend age twelve. In selecting ten he was conforming with the practical necessities of the day. Most children entered the factory as workers at this age. Post-elementary schooling would consist of two-hour night classes for children and tri-weekly sessions for adults. The children were to extend their knowledge of reading, writing, accountancy and for the girls sewing. The adults were to learn child care and training.



Owen recognized the importance of the proper training of teachers to qualify them for the very important responsibility assigned to them. He recommended that legislation be passed which would provide

For the establishment of seminaries in which those individuals who shall be destined to form the minds and bodies of the future subjects of these realms should be well initiated in the art and matter of instruction . . . Let the instruction to the young be well-devised and well executed, and no subsequent proceedings in the State can be materially injurious.<sup>17</sup>

The specific aim of education was to develop "full-formed men and women physically and mentally, who would always think and act consistently and rationally."<sup>18</sup> What would Owen have taught to promote this aim? Very important would be the inculcation of good habits and sentiments. Since any group of habits and sentiments can be developed it is most important that the proper ones be selected for development, those that will contribute to happiness. Basic here is his precept of never injuring one's fellows, and, on the contrary, of contributing to their happiness in every possible way. This, he maintained, represented rational conduct and would be reinforced in younger children by the example of the older children who would have come to understand and appreciate the merit of such conduct. Such behavior would develop into habit if no deviation was permitted.

Academically he would stress the mastery of basic skills and the development of a broad background of knowledge on which sound judgements could be based. Above all, it was important that the basic skills were applied to proper and useful knowledge and not to vicious knowledge.

With respect to physical education considerable emphasis was placed on dancing and training in military formations. Both were commenced at a very early age and both were expected to improve health, grace and strength, and the last was expected to prepare future citizens for the defence of the nation. Defence would not be necessary if all men were trained in a rational manner but until such was the case military preparedness was a necessity. He did think that it was much preferable to have military training take place in the school than to expose youth to the recruitment and training policies of his day. And from a national point of view he considered that a school training program would, in the course of time, provide a permanent armed force for which little national expenditure would be required.

In connection with health Owen gave his attention to dress. He considered that the Romans and the Scots had developed forms of dress best suited to promote health, grace and strength. He recom-

mended a combination toga and kilt type garment that would provide the freedom conducive to the development of these attributes.

Regarding methodology Owen was well attuned to the principles advocated by many educational reformers. He agreed that experience was the basis of knowledge and he maintained that the opportunities for using first-hand experience should be completely exhausted before resorting to second-hand experience. Field trips were considered very valuable. When first-hand experience of a primary sort was impossible secondary sources such as pictures and maps should be generously used.

The children in these new schools, he writes, should therefore be trained systematically to acquire useful knowledge through the means of sensible signs, by which their powers of reflection and judgement may be habituated to draw accurate conclusions from the facts presented to them.<sup>19</sup>

It was also important that learning should take place in a context. Only in this way would there be understanding. Most schools encouraged only rote memory which was a mockery of learning.

Thus the child . . . if . . . he possess a memory to retain incongruities without connection, will become what is termed the first scholar in the class; and three fourths of the time which ought to be devoted to the acquirements of useful instruction, will be really occupied in destroying the mental powers of children.<sup>20</sup>

Like Lancaster and Bell, Owen did not consider large classes to be a problem. Proper supervision over a large group was superior to the effects of too much attention that comes with two or three infants being placed under the care of some person who does not understand human nature.

Owen considered competition among students to be bad. In part this attitude stemmed from his own experience as a child when he was pitted against his brother in a writing contest and won. With respect to this incident he felt that he lost some of his brother's affection unnecessarily and he wrote:

I have said that such competitions are unjust, because, as no two organizations are the same, there can be no just comparison between the competing efforts of any two individuals,—while the successful one is thus taught vanity, and the unsuccessful jealousy and hatred.<sup>21</sup>

Closely associated with his view on competition was his view on discipline. Discipline should be very mild. Basically this was reasonable because since characters were made for and not by individuals they could not be held responsible for their behavior. Rather the teacher or society should be the object of punishment. In any case it was not the child who should be disciplined. He

Considered all children as beings whose dispositions, habits and sentiments are to be formed for them; that these can be well formed only by excluding all notions of reward, punishment and emulation; and that, if their characters are not such as they ought to be, the error proceeds from their instructors and the other circumstances which surround them.<sup>22</sup>



For immediate disciplinary results the best method is to keep children interested and busy—the worst is the use of the birch. He agreed with Quintilian and many others that harsh discipline could very easily kill all interest in learning for the young.

It has been and ever will be found far easier to lead mankind to virtue, or to rational conduct, by providing them with well-regulated innocent amusements and recreations, than by forcing them to submit to the useless restraints, which tend only to create distrust, and often to connect such feelings even with that which is excellent in itself, merely because it has been judiciously associated.<sup>23</sup>

But, in the long run, the only sound disciplining of human beings comes through the molding of good character and this, Owen was convinced, could be done by providing children with the sort of environment that would instil sentiments of charity from the day of their birth. Education in the sense of character training, as opposed to the idea of book learning, could not begin too soon in life. The infant school was a primary answer to many of society's most pressing problems.

## Part VI

It was in conjunction with the Scottish mill at Lanark that Owen first put his educational ideas into practice in 1816. In 1814 the third business partnership he had was formed, and these men were more agreeable to social experimentation than previous partners had been. He tells us that children who could walk were received into the school and some of these were under two years of age. Although he favored boarding schools the Lanark institution was a day school. He charged three shillings a year and he estimated that the cost per child was two pounds a year. He considered that the deficit was more than made up by the improved character of the population which the school brought about. The principles he had enunciated prior to the opening of the school were practiced. Education by 'sensible signs' and familiar conversation was employed. Books, he suggested, should not be employed before the age of ten. Field trips, maps and pictures were employed in a liberal manner. Kindness to their fellows was impressed upon the children. The older children between the ages of four and six were taught to be concerned about and to assist the younger ones.

The subjects taught were reading, writing, arithmetic, dancing, singing, military drill, botany, history and geography. His Quaker partners objected to the singing, dancing and military drill and in 1824 these subjects ceased to be a regular part of the program although dancing continued to be taught as an extra subject for which there was a special charge. This type of training was not to be a charge against the business. The partners also took exception to what they considered to be the immodesty of the dress or uniforms.



From the point of view of public appeal the banned subjects were among the ones that had interested visitors the most keenly.

Singing lessons began at the age of four. In the class there would be about one hundred fifty voices, which were trained to harmonize. It seems that the ability of the children to sing old Scottish songs was notable. The dancing also involved a large number of children at one time. Usually there were about seventy dancing couples on the floor at once and their skill was such that the presence of the dancing teacher was scarcely felt. The children began their dancing lessons as soon as they entered the school. The military formations were carried out by both boys and girls who went through their exercises with six or eight little pipers in the lead. The children's proficiency in executing their manoeuvres drew very favorable comment from the army officers who saw them.

It was, however, the geography lessons that created the greatest interest among visitors. At first these were made very specific lessons in tolerance as was Owen's general aim in all instruction. He hoped that the children could be taught that differences among people resulted from differences in environment and not from the operation of individual wills. The moral aspects of the geography lessons had to be dropped at the insistence of the partners. However, the lessons had other points to interest the visitors. Owen described the lessons as follows:

At a very early age they (the children) were instructed in classes on the maps of the four quarters of the world, and after becoming expert in a knowledge of these, all the classes were united in one large class and lecture room, to go through these exercises on a map of the world so large as to almost cover the end of the room. On this map were delineated the usual divisions of the best maps, except there were no names of countries or cities or towns; but for the cities and towns were small but distinct circles to denote their places—the classes united for this purpose generally consisted of about one hundred and fifty, forming as large a circle as could be placed to see the map. A light wand was provided, sufficiently long to point to the highest part of the map by the youngest child. The lesson commenced by one of the children taking the wand to point with. Then one of them would ask him to point to such a district, place, island, city, or town. This would be done generally many times in succession; but when the holder of the wand was at fault, and could not point to the place asked for, he had to resign the wand to his questioner, who had to go through the same process.<sup>24</sup>

The writer knows that many in his own times will have experienced similar geography lessons but not in the presence of so many fellow pupils.

The physical plant in which these lessons were conducted might be of interest. The following is the description Owen's son has left us of the New Lanark five roomed school.

The principle school-room, is fitted up with desks and forms, on the Lancastrian plan, having a free passage down the centre of the room. It is

surrounded, except at one end, where a pulpit stands, with galleries, which are convenient when this room is used, as it frequently is, either as a lecture-room or a place of worship.

The other and smaller apartment on the same floor has the walls hung round with representations of the most striking zoological and mineralogical specimens, including quadrupeds, birds, fishes, reptiles, insects, shells, minerals etc. At one end there is a gallery, adapted for the purpose of an orchestra, and at the other end are hung very large representations of the two hemispheres; each separate country, as well as the various seas, islands, etc., being differently coloured, but without any names attached to them . . . and it is here that the dancing and singing lessons are daily given. It is occasionally used as a reading room for some classes.

The lower story is divided into three apartments of nearly equal dimensions . . . It is in these three apartments that the younger children are taught reading, natural history, and geography.<sup>25</sup>

As evidence of the extremely favorable impression Owen's experiment made on visitors we can look at two assessments. A Leeds delegation reported that they found three hundred and eighty pupils between the ages of two and ten years of age. The report of the delegation reads, in part:

In the education of children the thing that is most remarkable is the general spirit of kindness and affection which is shown towards them, and the entire absence of everything that is likely to give them bad habits, with the presence of whatever is calculated to inspire them with good ones; the consequence is, that they appear like one well-regulated family, united together by the ties of closest affection.<sup>26</sup>

The physician to the Duke of Kent, Dr. Macnab, wrote:

The children and youth in this delightful colony are superior in point of conduct and character to all the children and youth I have ever seen.<sup>27</sup>

Owen considered that there were ten features of the New Lanark school that were new in educational practice. We are not likely to agree that his school was this original, but his points do indicate the degree to which he was in step with other educational reformers. The first two points stress the lack of traditional disciplinary measures and the emphasis on never failing kindness in the treatment the children received. Points three and four dealt with teaching through a study of the 'thing' and teaching through familiar conversation between the teacher and the pupil. Questions were invited and gaps in the teachers' knowledge were to be candidly acknowledged so as never to tamper with fact or truth. Point number five dealt with the informality of the school routine. There were no regular school hours. When the mind was keen and receptive teaching would take place and when the child was tired he was free to take part in exercises or other recreational activities such as music. Points six and seven refer to the character training derived from instruction in music, dancing and military exercises providing they were not overdone. Point number eight deals with field trips being practical and an essential feature of instruction at New Lanark. Points nine and ten maintained that



It was quite new to train children of the working class to think and act rationally, and to acquire substantial knowledge which might be useful to them through after life . . . It was quite new to place the child of the working man within surroundings superior to those of the children of any class . . .<sup>28</sup>

Robert Owen, one of the most altruistic and sincere of men, deserved much more recognition than he has ever received for his devotion to and concern for the genuine welfare of little children, that segment of humanity that is so often ill-treated either negatively or positively or both.

#### REFERENCES

1. Robert Owen, *The Life of Robert Owen*, G. Bell and Sons, London, 1920, pp. 3-4.
2. *Ibid.*, p. 49.
3. *Ibid.*, p. 122.
4. *Ibid.*, p. 116.
5. *Ibid.*, p. 185.
6. H. de B. Gibbins, *Industry in England*, Methuen & Co., London, 1915, p. 398.
7. Frank Smith, *A History of English Elementary Education, 1760-1902*, University Press, London, 1930, p. 40.
8. *Ibid.*, p. 41.
9. *Ibid.*, p. 48.
10. G. D. H. Cole, "Introduction," *A New View of Society and Other Writings*, by Robert Owen, J. M. Dent & Sons, London, 1927, p. v.
11. Robert Owen, "Catechism of the New View of Society and Three Addresses," *A New View of Society and Other Writings*, J. M. Dent & Sons, London, 1927, p. 216.
12. Robert Owen, *The Life of Robert Owen*, p. 181.
13. Robert Owen, "New View of Society, Third Essay," *A New View of Society and Other Writings*, p. 54.
14. *Ibid.*
15. *Ibid.*, Fourth Essay, p. 70.
16. *Ibid.*, First Essay, p. 16.
17. *Ibid.*, Fourth Essay, p. 82.
18. Robert Owen, *The Life of Robert Owen*, p. 185.
19. Robert Owen, "Report to the County of Lanark," *A New View of Society and Other Writings*, pp. 282-283.
20. Robert Owen, "New View of Society, Fourth Essay," *A New View of Society and Other Writings*, p. 76.
21. Robert Owen, *The Life of Robert Owen*, p. 6.
22. Robert Owen, "Report to the County of Lanark," *A New View of Society and Other Writings*, p. 282.
23. Robert Owen, "New View of Society, Third Essay," *A New View of Society and Other Writings*, pp. 43-44.
24. Robert Owen, *The Life of Robert Owen*, p. 199.
25. Frank Podmore, *Robert Owen*, George Allen & Unwin, London, 1923, p. 87.
26. *Ibid.*, p. 149.
27. *Ibid.*
28. Robert Owen, *The Life of Robert Owen*, pp. 320-321.



## BOOKS RECEIVED

*The Second Canadian Conference on Education, 1962.* University of Toronto Press, Toronto. Pp. 404, \$6.00.

The aims of the Conference were: (1) to improve communication among segments of Canadian society interested in education, (2) to help create wide public understanding and support among Canadians for educational development, (3) to encourage appropriate efforts designed to solve the problems, (4) to keep informed about emerging needs of education, and (5) to offer, as appropriate, to cooperate in the promotion of activities designed to arouse public interest in education.

*Townsend, E. A. and Burke, P. J. Learning for Teachers.* Collier MacMillan, New York, 1962. Pp. 313.

"This book is intended to bridge the gap between the general principles of psychology and the courses in methods of teaching. The concentration here is on the behaviour underlying teaching and learning." Typical of the eighteen chapter headings are those on Remembering, Transfer, Cognition and Communication.

*Bigge, Morris L. and Hunt, Maurice P. Psychological Foundations of Education.* Harper and Brothers, New York, 1962. Pp. 509, \$7.50.

A basic text in its field, dealing with all conventional topics in educational psychology and focused upon the nature of learning and its meaning for teaching. Although the authors acknowledge their preference for a broadly configurational point of view they deal fairly and objectively with major alternative views, particularly the most active one—psychological connectionism. Believing that the "cognitive field" outlook has more to offer teachers than any other they continually point out what this frame of reference means for teaching.

*Gabriel, John. Children Growing Up. The Development of Children's Personalities.* Clark, Irwin and Company, 1964. Pp. 367, \$6.30.

The author says: "The major part of this book is concerned in a descriptive and interpretive way with the meaning of personality, with the social and individual factors in its formation, and with the general patterns of behaviour which occur during the various stages of children's growth."

Its four parts, subdivided into chapters, are denominated: The Meaning of Personality; The Formation of Personality and its Organization; Children's Emotional and Social Development; and The Observational Study of Children.

# CONTENTS

Volume X, 1964

No. 1, March

	PAGE
Occupational Aspirations of Northern Saskatchewan Students ..... <i>William D. Knill</i>	3
Intelligence Tests for Two Samples of Metis and Indian Children ..... <i>L. W. West and R. S. MacArthur</i>	17
Learning Effectiveness Under the Trimester System ..... <i>D. A. Girard and F. Enns</i>	28
Textbook Selection for Grade Ten Language ..... <i>James B. Bell</i>	34
Problems Inherent in Ability Grouping ..... <i>Arthur A. Attwell and Thomas E. Linton</i>	39
Leisure Reading in the Senior High Schools of Alberta ..... <i>C. S. H. Campbell</i>	46

No. 2, June

Comparison of Matched Groups in Manchester, England and Edmonton Alberta ..... <i>J. A. Young and M. D. Jenkinson</i>	59
A Comparative Study of C. P. Snow's "Two Culture" Controversy ..... <i>Donald F. Swift and Thomas E. Linton</i>	67
A Study of Student Assessment of Teacher Performance ..... <i>B. E. J. McBride</i>	79
Organization for Teaching Physical Education in Elementary Schools ..... <i>S. Boyd Taylor and Walter H. Worth</i>	90
Prediction Validity of the Metropolitan Readiness Tests ..... <i>Harvey W. Zingle and A. E. Hohol</i>	99
Textbook Difficulty and Mathematics Achievement in Junior High School ..... <i>L. Doyal Nelson</i>	105
The Human Factor in Critical Analysis ..... <i>Norman Cuthbertson</i>	114

### No. 3, September

	PAGE
Validity of Regression Equations After Six Years .....	125
<i>D. B. Black</i>	
The Knowledge of Mathematics of Alberta Students .....	137
<i>Ellen A. Laws</i>	
Reference Group Behaviour Among High School Students .....	142
<i>Henry Zentner</i>	
The Improvement of Methods Courses and Practice Teaching .....	153
<i>Merron Chorney and Arthur G. Storey</i>	
Incongruity Tolerance and the Counselling Setting .....	158
<i>G. B. Jones and C. M. Christensen</i>	
Stereotyping and the Use of All-Inclusive Terms .....	170
<i>James G. Snider</i>	
Book Review .....	175
<i>W. D. Neal</i>	

### No. 4, December

Sir Richard Livingstone on Education .....	181
<i>Barbara Hutchinson</i>	
The Academic Interests of Failing College Students .....	188
<i>E. R. Oetting</i>	
Differential Achievement in Reading and Arithmetic .....	192
<i>Mary H. Neville and Barry P. Frost</i>	
An Investigation to Identify Creativity in Teaching .....	201
<i>David J. Chabassol and Gordon G. Manson</i>	
Some Effects of Mental Health Instruction .....	209
<i>Metro Gulutsan</i>	
Relationship between Achievement Scores and Self-estimates .....	217
<i>S. B. Kakkar</i>	
A Survey of Literature Concerning the Place of Grammar .....	220
<i>Howard G. Ambury</i>	
Yesterday's Infant School .....	234
<i>K. H. Thomson</i>	











